Adaptive Human-Computer Interfaces Using Expert Profiles

Robert C. Williges, Jay Elkerton, Kim J. Vicente,
and Brian C. Hayes

Virginia Polytechnic Institute and State University

for

Contracting Officer's Representative
Michael Drillings

DTIC
ELECTE
SEP 27 1990
S B D

Basic Research
Michael Kaplan, Director

July 1990

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

NOTICES

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	--

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
--	Approved for public release;
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	distribution is unlimited.
--	

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
HCI-87-01	ARI Research Note 90-83

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Virginia Polytechnic Institute and State University	--	U.S. Army Research Institute

6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
130 Whittemore Hall Virginia Tech Blacksburg, VA 24061	5001 Eisenhower Avenue Alexandria, VA 22333-5600

8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Army Research Institute for the Behavioral and Social Sciences	8b. OFFICE SYMBOL (If applicable) PERI-BR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER MDA903-84-C-2017

8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
5001 Eisenhower Avenue Alexandria, VA 22333-5600	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	61102B	74F	N/A	N/A

11. TITLE (Include Security Classification)

Adaptive Human-Computer Interfaces Using Expert Profiles

12. PERSONAL AUTHOR(S)

Williges, Robert C.; Elkerton, Jay; Vicente, Kim J.; and Hayes, Brian C.

13a. TYPE OF REPORT Final	13b. TIME COVERED FROM 86/09 TO 87/09	14. DATE OF REPORT (Year, Month, Day) 1990, July	15. PAGE COUNT 131

16. SUPPLEMENTARY NOTATION

Contracting Officer's Representative, Michael Drillings

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Human factors Cognitive psychology
			Computer science Human-computer interaction (KR)(
			Educational psychology Novice (Continued)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Adaptive human-computer systems accommodate a wide variety of users learning to interact with computers because they adjust to different skill levels and provide novices with appropriate levels of expertise needed to perform certain tasks. This effort was directed toward developing improved models of experts based on goal-based models and toward assaying and isolating individual differences of inexperienced users in order to adapt the software interface to these individual differences.

Results show that inexperienced computer users were more variable in their search times than experienced users and that slower users selected more inefficient search commands. Two performance-based and one cognitive-based command selection aid improved search performance and strategies of slower, inexperienced users Since spatial and verbal ability were found to correlate positively with search strategies, inexperienced users learned to select fewer and more efficient commands when provided with spatial augmentation (graphic presentation).

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Michael Drillings	(202) 274-8722	PERI-BR

DD Form 1473, JUN 86   *Previous editions are obsolete.*   SECURITY CLASSIFICATION OF THIS PAGE

i

ARI Research Note 90-83

18.   SUBJECT TERMS (Continued)

Adaptive
Learning strategies
Expert
Computer-based instruction
Training
Machine learning

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

ACKNOWLEDGMENTS

ADAPTIVE HUMAN-COMPUTER INTERFACES USING EXPERT PROFILES

CONTENTS

ADAPTIVE HUMAN-COMPUTER INTERFACES USING EXPERT PROFILES

RESEARCH PROGRAM

Background

Computer-based systems are being used by people with various levels of experience. Consideration needs to be given to develop ways to accommodate these differences if computers are to become truly useful tools. In particular, inexperienced users of computer-systems often become overwhelmed and frustrated by complex computer interfaces thereby causing them to be unable to complete, or extremely inefficient in completing, computer-based tasks. One way to accommodate these user differences is to adapt the software interface according to the capabilities and limitations of the users.

Human/computer communication occurs primarily as a static interface on most computer-based systems. The computer usually responds to the user in a predefined manner, and specified human-computer dialogue design principles are used to optimize a particular interface. Such a static approach to design may not accommodate all interface situations. An alternative to this approach is to provide an adaptive interface in which the dialogue is tailored to the characteristics of the user, the level of experience of the user, and the task configuration. Such an approach has implications for expert aiding in providing online assistance for novice users of computer systems. To provide this adaptive assistance, one must combine the knowledge of experts into an adaptive interface.

Adaptive human-computer systems could provide several advantages over existing static interfaces. First, the prevalence of computer systems is constantly increasing. Therefore, the number of people who work with computers is also growing. Providing a static interface that accommodates the many types of users is an impossible task. A computer systems that can adapt to the wide range of users is almost a necessity. Second, it is no longer reasonable to expect users to be

1

familiar with the computer system. Users rightfully expect the computer system to be easy to use regardless of the user's level of expertise. An adaptive computer system is, by design, capable of adjusting to different skill levels. And, third, adaptive interfaces can provide a convenient delivery mechanism for providing the appropriate level of expertise to aid the novice user in performing a specific task.

Research Objectives

A three year research program was conducted to investigate critical issues in the design of an adaptive human/computer interface using on-line expert profile generation procedures. The general objective of this program was to gain an understanding of how to capture and use expert knowledge to enhance human/computer interfaces. Specifically, this research refined the expert profile methodology, developed procedures for adaptive model building, and incorporated expertise into an adaptive assistant for aiding novice users of computer systems. The results of this research provided an understanding of critical components of intelligent interfaces which are tailored to individual users of complex, computer-based systems.

Approach

A series of related research studies was conducted during this research program. These studies all dealt with adapting the human-computer interface to aid inexperienced users in locating information contained in a hierarchical file system.

Search task environment. The hierarchical file system was contained in a task environment that was implemented on a VAX 11/750 under the VMS operating system. Two Digital VT100 terminals were used for these studies. One display was used for information retrieval, and the other was used to present the targets. As

2

shown in Figure 1, the information retrieval display contained a 7-line window, a line containing the name and length of the current file, an input and message line, a work area, and a touch keypad. Commands were selected through a Carroll Touch Technology touch-entry device and a VT100 (Qwerty) keyboard.

The information retrieval task required the subjects to retrieve a single line of information — the target — from a 2,780 line hierarchical file system developed by Elkerton (1985). This file system consisted of 15 files ranging in length from 55 to 447 lines. Each file contained hypothetical textual and numerical information pertaining to army operations, military vehicles, or combat support.

Interactive search commands. The touch keypad allowed subjects to select from any of the 12 search commands shown in Figure 1. These commands included file manipulation commands, large movement search commands, and small movement commands. The file manipulation commands allowed the subjects to move throughout the file hierarchy, whereas all other commands were restricted to the information contained in the current file. The FILE SELECT command was used for direct selection of a lower-level file. Within the top level files in the hierarchy, certain lines were highlighted to indicate that additional information on these selected topics was available in other files. The ZOOM IN command could be used whenever a highlighted line appeared in the window to move to a lower level file in the hierarchy. Conversely, the ZOOM OUT command allowed the subject to exit from the current file and position the top level file in the window.

The large movement commands either allowed for direct string searches (SEARCH, SEARCH-AND SEARCH-NOT) or exploited the structure of the current file (SECTION and INDEX). The string search commands allowed the subjects to locate specific alphanumeric strings in the current file. The SEARCH command permitted a

3

Figure 1. Information retrieval display used in the search task environment.

simple string search, and the SEARCH-AND and SEARCH-NOT commands allowed multiple string searches in the 7-line window using Boolean operators. The SECTION command is analogous to a table of contents in a book, in that it presented a list of the major topics contained in the current file. The window could be positioned to a specific location in the current file by choosing the appropriate topic. The INDEX command presented the subject with an alphabetized list, in the work area of the display, of keywords contained in the current file. The window was positioned to that area of the file containing the chosen keyword.

The small movement commands allowed the subject to scan the current file either seven lines at a time (PAGE UP and PAGE DOWN), or one line at a time (SCROLL UP and SCROLL DOWN).

Scientific Reporting

Presentations. Throughout the three-year research program a variety of scientific presentations and reports were prepared. Over twenty presentations related to this research effort were made at professional society meetings, colloquia, and symposia. These included technical presentations made at:

American Psychological Association Meeting

Ergonomics Society Meeting

Human Factors Society Meeting

IEEE International Conference on Systems, Man, and Cybernetics

IFIP Conference on Human-Computer Interaction

Interactive Computer Systems Symposium

International Conference on Human Aspects of Computing

Third Midcentral Ergonomics/Human Factors Conference

Digital Equipment Corporation

IBM Watson Laboratory

RISO National Laboratory

George Mason University

Georgia Institute of Technology

New Mexico State University

Rice University

The Ohio State University

University of California, Los Angeles

University of Illinois

University of Michigan

Graduate degrees. In addition to these technical presentations, a variety of scientific publications were completed to document the research completed on this program. Some of the specific research issues addressed in this research program resulted in a graduate thesis and dissertation:

> Elkerton, J. (1985, December) A behavioral evaluation of command-selection aids for inexperienced computer users, Ph.D. Dissertation, Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University.

> Vicente, K.J. (1987, March) Accommodating individual differences in searching a hierarchical file system, M.S. Thesis, Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University.

Archival publications. Fourteen archival publications were completed on this project. These publications include a book, book chapters, journal articles, and papers printed in the proceedings of scientific meetings. The complete citation of each of these publications is as follows:

Elkerton, J. (1985, October) Modeling expert computer users to aid inexperienced computer users. *Proceedings of the Human Factors Society 28th Annual Meeting* (pp. 851-855) Santa Monica: Human Factors Society.

Ehrich, R.W. and Williges, R.C. (Eds.) (1986) *Human-Computer dialogue design.* Amsterdam: Elsevier

Hayes, B.C. and Williges, R.C. (1986, October) Defining search strategies in information retrieval. *Proceedings of the 1986 IEEE International Conference on System, Man, and Cybernetics,* (pp. 1092-1096) Ne / York: IEEE

Hayes, B.C., Vicente, K.J. and Williges, R.C. (1986, October) Command selection differences among users in an information retrieval task. *Proceedings of the 1986 IEEE International Conference on System, Man, and Cybernetics,* (pp. 1108-1112). New York: IEEE

Williges, R.C., Williges, B.H., and Elkerton, J. (1987) Software interface design. In G. Salvendy (Ed.) *Handbook of human factors* (pp. 1416-1449). New York: Wiley.

Vicente, K.J., Hayes, B.C., and Williges, R.C. (1987) Individual differences in computer-based information retrieval. In L.S. Marks, J.S. Warm, and R.L. Huston (Eds.) *Ergonomics and human factors* (pp. 225-229). New York: Springer-Verlag.

Williges, R.C. (1987) The use of models in human-computer interface design. *Ergonomics,* **30,** 491-502.

Vicente, K.J., Hayes, B.C., and Williges, R.C. (1987). Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors,* **29,** 349-359.

Elkerton, J. (1987) A framework for designing intelligent human-computer dialogues. In G. Salvendy (Ed.) *Cognitive engineering in the design of human-computer interaction and expert systems*, (pp. 567-574), Amsterdam: Elsevier.

Williges, R.C. (1987) Adapting human-computer interfaces for inexperienced users. In G. Salvendy (Ed.) *Cognitive engineering in the design of human-computer interaction and expert systems*, (pp. 21-28), Amsterdam: Elsevier.

Elkerton, J. and Williges, R.C. (1987, September) A summary of experimental research on command-selection aids. *Proceedings of the Second IFIP Conference on Human-Computer Interaction* (pp. 937-942), Amsterdam: Elsevier.

Vicente, K.J. and Williges, R.C. (in press) Visual momentum as a means of accommodating individual differences among users of a hierarchical file system. *Proceedings of the Fifth Symposium on Empirical Foundations of Information and Software Science*. Denmark: Riso National Laboratory.

Elkerton, J. and Williges, R.C. (in press) Dialogue design for intelligent interfaces. In M.H. Chignell, P.A. Hancock, and A. Loewenthal (Eds.) *Intelligent interfaces: theory, research, and design*. Amsterdam: Elsevier.

Vicente, K.J. and Williges, R.C. (submitted) Accommodating individual differences in searching a hierarchical file system. *International Journal of Man-Machine Studies*

SUMMARY OF RESEARCH FINDINGS

This section provides a brief summary of the research findings over the integrated, three-year research program. Essentially, the overall scientific effort was directed toward research to refine expert profile methodolgy, to develop procedures for adaptive model building, and to provide online adaptive assistance in computer tasks. This summary is divided into three major sections which show the progression of the research throughout the three-year program. Appendices A, B, C, and D, provide a more detailed description of the research procedures, findings, and discussion summarized in this section. Specifically, Appendix A provides details on the "Aiding Search Performance" section. Appendix B provides a detailed discussion of the "Goal-Based Models" section. And, Appendices C and D provide the details related to the "Individual Differences" section.

Aiding Search Performance

Command-selection models. Elkerton (1985) demonstrated marked differences between 12 highly experienced and 12 inexperienced users in retrieving information from the hierarchical file system. These differences manifested themselves in terms of search times, total number of search commands, and choice of search commands. In particular, inexperienced computer users were more variable in their search times than experienced users. Additional analyses on the slower inexperienced users data demonstrated that these slower users selected the more inefficient search commands of SCROLL UP, SCROLL DOWN, and ZOOM OUT than their counterpart faster inexperienced users. These faster users more closely resembled the command selection profiles of the experienced computer users.

Given these marked differences between inexperienced and experienced users, models were required which inexperienced users could use to improve their search performance. Two performance-based models and one cognitive plan-based model were developed to explore a variety of representations for command-selection information.

The first performance-based model was a frequency profile of commands used by experts based on a polling procedure methodology developed by Elkerton and Williges (1985). The second performance-based model added additional data on command sequence in terms of a second-order transition matrix to the basic command frequency profile model. The third, cognitive model extended the frequency and sequence models with detailed search plans which provided ordered lists of search procedures and supporting explanations for using search commands.

Evaluation of command-selection aids. The three models were implemented as online aids for inexperienced users and evaluated with a new sample of 36 inexperienced users. These command-selection aids improved search performance and strategies of the slower inexperienced users when the underlying command-selection models did not involve complicated second-order transition matrix sequence information and did not introduce a highly interactive, mixed-initiative dialogue which allowed online aiding by either user or computer initiation. These results were successful in demonstrating a method for providing inexperienced computer users with online assistance and instruction based on the expert user's search performance profiles. Two unresolved issues remained. First, there was a requirement to investigate ways to improve the model of the expert user's search strategy. Second, Elkerton (1985) found marked individual differences among the inexperienced users which suggested that further tailoring of the adaptive

10

interface may be required. Both of these issues were addressed in follow-on research.

Goal-Based Models

Model development. In reviewing the Elkerton (1985) data and in considering the hierarchical nature of the computer files, Hayes and Williges (1986) noted that the overall goal of identifying a target line could be decomposed into two subgoals. The first subgoal was to locate the appropriate file. The second subgoal was to locate the appropriate line within the file.

A subsequent task analysis of the 96 targets used by Elkerton (1985) showed that his targets appeared to cluster into three distinct groups. The first group was designated an explicit target group, because the description of the target revealed the file in which the target line appeared. The second group of targets was defined as the clued group, because those target descriptions provided no file identity information. And, the third group of targets was designated a non-clued group, because those target descriptions provided no file identity information.

Each target group was analyzed according to the task goal and subgoals according to the GOMS model of Card, Moran, and Newell (1983) to develop a model of experienced users' search strategies. These models are depicted in terms of a flow diagram which represent the list of search commands and their sequence of use for each target group.

Model validation. Data were collected by Hayes and Williges (1986) on a new group of 12 experienced computer users to validate these models. Each of the expert users searched for 108 target lines, 36 targets for each of the three target types. The order of presentation was randomized, and the users had no knowledge of the existence of the three target types.

11

Two performance measures were used to assess the predictive validity of these models. First, search performance in terms of time to locate the appropriate file line and the total number of search commands confirmed that the explicit target grouping was the least difficult, followed by the clued and non-clued target groups as the models suggest. Second, the relative proportion of use of the twelve possible search commands in each target grouping had a remarkably close resemblance to the search commands in three models. Namely, the commands used in the models were the highest proportion of commands actually used in each target grouping.

Aiding inexperienced users. The three validated models were used as a means of aiding inexperienced users in searching for target lines. An aiding system was implemented into the software interface that provided advice to the inexperienced users when their search strategy did not follow the goal-based model of experienced users' performance. The advice was presented in the form of a suggestion to follow the search commands represented by the three experienced users' models.

A group of 12 inexperienced users searched for 36 target lines while receiving this advice and then transferred to a task which required searching for an additional 18 target lines with no advice available. The inexperienced users who received advice were compared to a control group of 12 inexperienced users who received no advice. The average of the last 18 practice trials and the average of the 18 transfer trials showed a significant improvement for the aided users. Clearly, the goal-based models were successful in improving inexperienced user performance in both practice and transfer.

12

Individual Differences

Elkerton (1985) also found that some of the inexperienced users learned to use the hierarchical filing task very quickly and actually exhibited target line retrieval performance equivalent to expert users. Other inexperienced users had extreme difficulty with the task. Consequently, these results suggest that individual differences may be an important component in tailoring the computer software interface. Egan and Gomez (1985) proposed a three part methodology to tailor the interface. First, an assay of the important sources of individual differences in task performance must be performed. Second, these individual differences must be isolated by identifying specific components of the task that are causing performance differences. And, third, the individual differences must be accommodated by modifying the isolated task components primarily responsible for the performance differences. Vicente, Hayes and Williges (1987) addressed the first two stages of this methodology and Vicente and Williges (1987) addressed the third stage. These results are summarized as an example of adapting software interfaces to accommodate individual differences.

Assaying individual differences. Vicente, Hayes, and Williges (1987) pretested a group of 30 inexperienced users on a battery of 21 predictors of individual differences including demographic measures, cognitive style, verbal abilities, and spatial abilities. Following this pretesting, each of the inexperienced users searched for target lines in the hierarchical filing system according to the procedures used in previous studies. Pretest measures were correlated with performance on the retrieval task, and the six significant predictors were entered into a stepwise regression to determine the best overall prediction equation. The resulting prediction equation ($R^2 = .45$) included two predictors as follows:

$$T = 516.29 - 1.67V - 16.09S$$

where T = the time to detect a target line, V = verbal ability as measured by the vocabulary section of the Nelson-Denny Reading Test (1973), and S = spatial ability as measured by the spatial visualization test from the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, and Harmon, 1976). Spatial ability was the more important of the two predictors in that it accounted for 33% of the variance as compared to 12% of the variance accounted for by the verbal ability predictor.

Isolating individual differences. A comparison of command selection strategies of subjects with high and low spatial ability was made to isolate these individual differences. The analysis of search command selection was based on a polling procedure developed by Elkerton and Williges (1985) which represented the relative use of each search command. The three significant search commands which distinguished high and low spatial ability users were ZOOM OUT, SCROLL UP, and SCROLL DOWN. In all cases, the low spatial ability users showed a high frequency of use of these commands. These commands, in turn, represent being lost in the hierarchical file structure and using inefficient search procedures. In fact, Elkerton (1985) observed the same pattern of differences between the detection times for slow inexperienced users and experienced users.

Accommodating Individual Differences. In order to accommodate the inefficient strategies that users with low spatial visualization abilities seem to be using, Vicente and Williges (1987) provided visual augmentation to the software interface. This visual augmentation was in the form of providing a graphical presentation of the hierarchical structure of the filing system as well as a visual indication of the user's relative position within the current file. The efficacy of the presence or absence of visual augmentation was compared factorially with high and low spatial ability on a new group of inexperienced users. The results supported the use of visual augmentation both in reducing the total number of commands and in

14

reducing an inefficient search command selection. Interestingly, the spatially augmented interface improved performance for both the high and low spatial ability users.

Some Unresolved Issues

Although these studies were successful in demonstrating the efficacy of aiding inexperienced users, several unresolved issues remain in order to improve adaptive interfaces.

Improved aiding. Both the goal-based models and the accommodation of individual differences investigated in this series of studies can be improved. The goal-based models need to be extended to other task domains in order to test the generalizability of this approach. Improved presentation methods need to be investigated in regard to these models. A deeper understanding is needed of the appropriate format for accommodating individual differences through graphical and verbal augmentation. The results of these studies are particularly encouraging for future investigations and development of workstations with direct manipulation interfaces.

Combined aiding. Goal-based models and individual differences were investigated separately. Further improvements in aiding inexperienced users may be realized by combining both of these approaches. For example, the display formats could be adapted based on individual differences and combined with aiding through goal-based models.

General strategies. In this series of studies, the interface was adapted in a highly proceduralized task as a means of aiding the inexperienced user. More generalized tasks and strategies need to be developed in order to provide an overall model of adaptive software interfaces. Such a model, in turn, may be useful as a tool

for designers of software interfaces (Williges, 1987). In addition, Williges, Williges, and Elkerton (1987) reviewed three strategies for developing adapting software interfaces. These included strategies to adapt the human/computer dialogue, strategies to adapt computer-based instructional systems which are embedded in the software interface, and strategies to adapt online assistance provided to the end-user. The current set of studies only addressed these strategies tangentially. Additional research is needed to focus the implications and integration of these strategies.

Conclusion

Based on the results of this research, it appears that adapting the software interface through goal-based models and individual differences can improve inexperienced users' performance. These results hold promise for a more extensive development of adaptive software interface models and concepts which will generalize across a variety of task domains. Such an approach can provide a truly user-centered software interface philosophy which would accommodate a wide variety of users.

References

Card, S. Moran, R.P., and Newell, A., (1983). *The psychology of human-computer interaction.* Erlbaum, New Jersey.

Egan, D.E., and Gomez, L.M., (1985). Assaying, isolating, and accommodating individual differences in learning a complex skill. In: R.F. Dillon (Editors), *Individual Differences in Cognition, Volume 2.* Academic Press, New York, pp. 173-217.

Ekstrom, R.B., French, J.W., and Harmon, H.H., (1976). *Manual for Kit of Factor-Referenced Cognitive Tests*. Educational Testing Service, Princeton.

Elkerton, J. (1985). A behavioral evaluation of command-selection aids for inexperienced computer users. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Elkerton, J. and Williges, R.C., (1985). A performance profile methodology for implementing assistance and instruction in computer-based tasks. *International Journal of Man-Machines Studies*, 23, 135-151.

Hayes, B.C. and Williges, R.C., (1986). Defining search strategies in information retrieval. In: *Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, New York, 1108-1112.

Nelson-Denny Reading Test, Form C, (1973), Houghton Mifflin, Boston.

Vicente, K.J., Hayes, B.C., and Williges, R.C., (1987). Assaying and isolating individual differences in searching a hierarchical filing system. *Human Factors*, 29, 349-359.

Vicente, K.J. and Williges, R.C., (1987). Accommodating individual differences in searching a hierarchical file system. Submitted to *International Journal of Man-Machine Systems*.

Williges, R.C., Williges B.H. and Elkerton, J., (1987). Software interface design. In: G. Salvendy (Editor) *Handbook of Human Factors*, Wiley, New York, 1416-1449.

Williges, R.C., (1987). The use of models in human-computer interface design, *Ergonomics*, 30, 491-502.

APPENDIX A

A SUMMARY OF EXPERIMENTAL RESEARCH ON

COMMAND-SELECTION AIDS[1]

Jay Elkerton

Center for Ergonomics, Department of Industrial and Operations Engineering

The University of Michigan, Ann Arbor, MI 48109

and

Robert C. Williges

Human-Computer Laboratory, Department of Industrial Engineering and Operations

Research, Virginia Polytechnic Institute and Sate University, Blacksburg, VA 24061

Two experiments were conducted to assess and improve novice performance in hierarchical file search. In the first experiment, half of the novices were performing like experienced subjects, using search procedures that exploited the structure of the file hierarchy for efficient information retrieval. Other slower novices used time-consuming scrolling strategies and appeared unfamiliar with the file hierarchy. In the second experiment, command-selection aids improved the search performance and strategies of slow novices when the underlying command-selection models were not unduly complicated (i.e., frequency profiles and search plans) and did not introduce a highly interactive, mixed-initiative dialogue.

---

[1] This paper appeared in the *Proceedings of the Second IFIP Conference on Human-Computer Interaction*

1. Introduction

Command selection is a basic skill required for proficiency with a computer interface. Learning a command language, however, is not an easy task. In addition to understanding the names and functions of commands, users must determine the appropriate strategies for using commands in a task. This research addressed the latter problem, the learning of conditions and procedures for command selection by users. From a study of users' search strategies, command-selection aids were designed and evaluated to help novices in hierarchical file search.

1.1 PRELIMINARY RESEARCH

Elkerton and Williges (1, 2) found large differences in the performance of novice and experienced subjects in a file search task. In these studies, experienced subjects selected powerful search procedures that used file content and structure for fast and efficient retrieval (e.g., string search and a procedure for accessing a list of file sections). In contrast, novices visually scanned and directly manipulated the files with time-consuming scrolling procedures.

Elkerton and Williges (3) implemented an online aid for these novices. Essentially, search commands selected frequently by experienced users were provided automatically to novices. This research revealed that presenting these search procedures could train novices to select more powerful commands. However, the method was not adequate to improve novice performance on the current information retrieval problem. This mixed effectiveness of the file search aid may be explained by the limitations of the frequency-based model and the constrained, computer-initiated dialogue. The frequency-based model provided no sequence, planning, or procedural assistance for command selection. In addition, the computer-initiated dialogue interrupted novices and provided them no opportunity to

query the file search aid. Nevertheless, the use of this impoverished dialogue for improving novice search skills suggested that further research determine how command-selection strategies be represented and communicated to inexperienced users.

1.2. RELATED RESEARCH ON LEARNING BY DOING

Research on command-selection aids also was motivated by investigations of learning by doing. Learning by doing is best exemplified by users actively learning a computer interface rather than reading documentation or following training materials (4). Interface support for learning by doing was studied initially by Carroll and Carrithers (5). These investigators found an effective learning environment in a "training wheels" interface which blocked advanced word-processing operations. The training wheels interface allowed users to explore the word processor without suffering the negative consequences of error states commonly associated with advanced features of the interface. In fact, Catrambone and Carroll (6) have shown recently that users were faster at learning more about the word processor after using the training wheels interface than users trained with the full interface. Thus, a training wheels approach may serve to decrease the complexity of the user interface and allow users to learn basic functions which ultimately can improve the acquisition of more advanced interface functions.

Unfortunately, reducing the complexity of the user interface to create a more effective learning environment is not an easy task. As an example, Carroll and Kay (7) have simplified the training wheels approach to the point of providing a single method or "scenario" for new users creating and printing a document. With each scenario, Carroll and Kay (7) manipulated the prompting, feedback, and error correction given to a user. The results indicated that "scenario machines" which

provide too much prompting and feedback, may destroy the coherence of the computing task and potentially eliminate any transfer of training. Thus, training dialogues must be designed carefully to call attention to important details without significantly altering the similarity of the interfaces in the training and transfer tasks.

2. Experiment 1: Development of the Command-Selection Models

Providing effective command-selection advice required the identification of novice difficulties in a hierarchical file search task. Therefore, the goals of this experiment were to identify novice-expert differences in search performance and to develop command-selection models which could help novices with their problems.

2.1 METHOD

2.1.1. Subjects. Twenty-four subjects participated in the experiment. Twelve were experienced subjects (experts) having used the computer for at least 2 years and twelve were inexperienced subjects (novices) having less than 20 hours of computing experience.

2.1.2. Experimental procedures. All subjects were familiarized with the information retrieval task and search procedures during the first two days of the six-day experiment. These subjects were screened at the end of the second day for a minimum proficiency with the file hierarchy and search procedures. On each of the subsequent four experimental days, subjects retrieved information on 24 search problems as quickly as possible.

2.1.3. Information retrieval task. Subjects retrieved information from 15 files arranged in a three-level hierarchy. These files contained both textual and numerical

data on a hypothetical army. Tables, hierarchies, lists, and paragraphs were used to structure information in each file. The hierarchical file system was implemented on two Digital VT100s connected to a dedicated VAX 11/750 computer. The primary VT100 displayed individual files in the hierarchy, while the secondary vt100 presented the search problems. The primary display contained a seven-line window, a file status line, an input and message line, a work area, and a touch keypad. The touch keypad was implemented with a Carroll Technology touch-entry device to facilitate a flexible display for search procedure selection.

2.1.4. Search procedures. Four search procedures were methods for small continuous movement within a file. The SCROLL UP and DOWN procedures moved the file a line at a time, while the PAGE UP and DOWN procedure presented consecutive seven-line segments of a file. Five additional search procedures were methods for large movement within a displayed file. The SECTION procedure presented a menu of file sections. Similarly, the INDEX procedure presented an alphabetical list of file topics. Finally, string search procedures included a simple string search (SEARCH) and multiple string searches with Boolean operators (SEARCH-AND, and SEARCH-AND-NOT).

The last three search procedures were file selection methods for traversal of the file hierarchy. The ZOOM IN and OUT procedure updated the file displayed in the seven-line window by moving up and down the file hierarchy, while the FILE SELECT procedure allowed users to select files directly from a menu in the work are of the primary display.

2.2. RESULTS AND DISCUSSION

2.2.1. Search performance. Using an analysis of variance (ANOVA), differences between novices and experts were confirmed over the four retrieval days for average search time ($F(1,22)$ = 8.59, $p < 0.01$). Compared to experts, novices took about 40% more time to find information. This relative difference was less than the previous 2:1 novice-expert performance ratios observed in other file search environments (1,2) prompting further analysis on the variability of search performance.

Differences in the variability of novice and expert groups were observed for average search time as illustrated by the box plots in Figure 1. In fact, a *F*- ratio of the variances for novices and experts (α^2 = 869.3 and α^2 = 80.4, respectively) showed significant differences in the variability of average times between these groups (F [11,11] = 10.8, p < 0.01). Further empirical comparisons revealed two possible novice groups with 6 subjects each: a slow group with search times approximately twice that of experts and a fast group with search times equal to experts. This heterogeneity of novice search performance resembles data presented by Schneider (8) on acquiring high-performance skills and suggests that command-selection aids should focus on improving the search skills of slow novices.

Insert Figure 1 Here

2.2.2. Search strategy. Results from a MANOVA (Wilks' λ [2,21] = 0.032, p < 0.01) and subsequent unbalanced ANOVAs on selection proportions for the 12 search procedures supported the hypothesis of diverse novice skills. Significant differences were found between slow novices, fast novices, and experts in the selection of
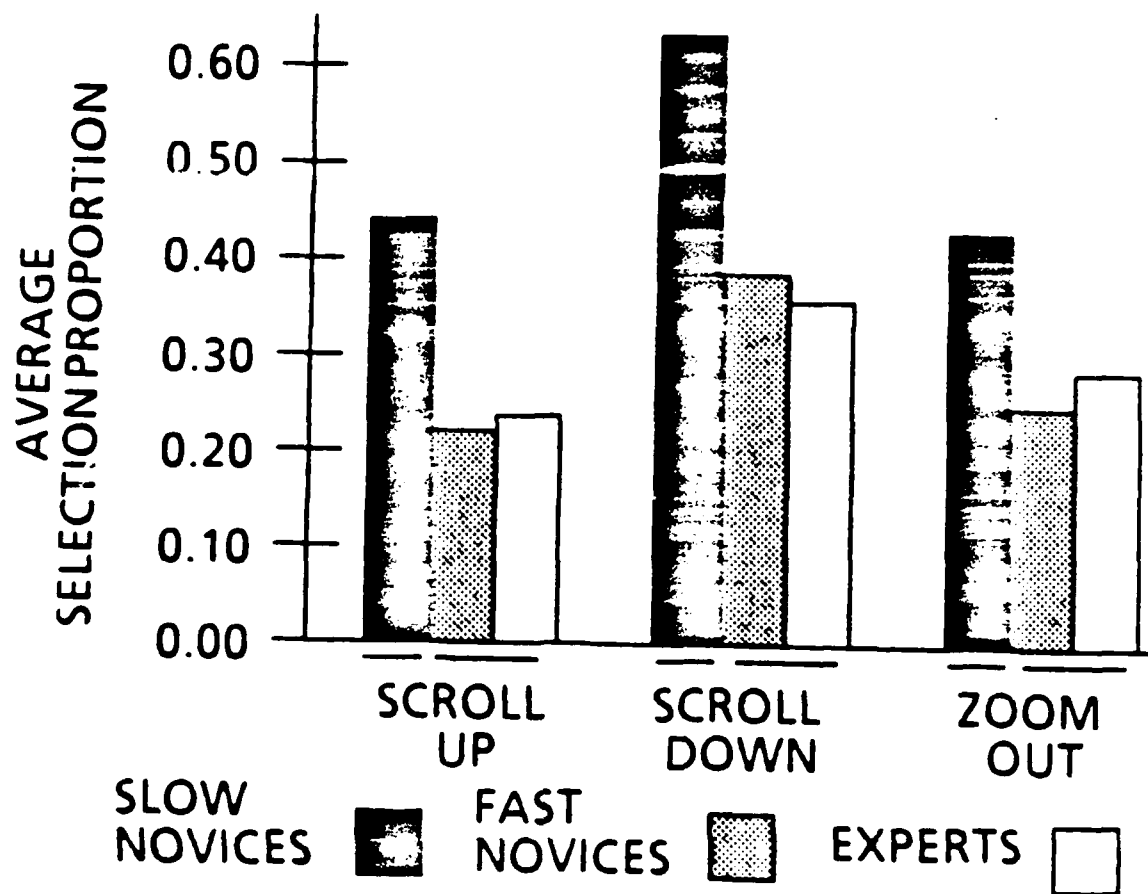
SCROLL UP (F (2,21) = 4.57, $p < 0.05$), SCROLL DOWN (F [2,21] = 5.85, $p < 0.01$), and ZOOM OUT (F [2,21] = 11.00, $p < 0.001$). Average selection proportions and the results of *post hoc* Newman-Keuls analyses ($\alpha = 0.05$) are shown for these search procedures in Figure 2.

The frequent selection of scrolling procedures by slow novices replicated the findings of Elkerton and Williges (1,2) indicating that slower, inexperienced subjects used inefficient scrolling procedures which directly manipulate files on the display. These scrolling strategies are time consuming and may not facilitate skill acquisition in the search task. To apply a training principle described by Schneider (8), excessive scrolling should be eliminated so that slow novices can spend more time learning other strategies rather than scanning the files visually.

The frequent selection of ZOOM OUT by slow novices suggested a lack of familiarity with the structure of the hierarchy. The ZOOM OUT procedure is used typically when the subject is in the wrong file. Therefore, an online aid should provide guidance on the location of the search problem in the file hierarchy.

2.2.3. *Command-selection models.* Given these novice-expert differences, models were required which novices could use to improve their search skills. Two performance-based models and a cognitive, plan-based model were developed to explore a variety of representations for command-selection information.

Insert Figure 2 Here

The first performance-based model, a profile of commands used frequently by experts, was developed and evaluated in previous investigations (2,3). The frequency-based method proved to be an easily constructed, but incomplete
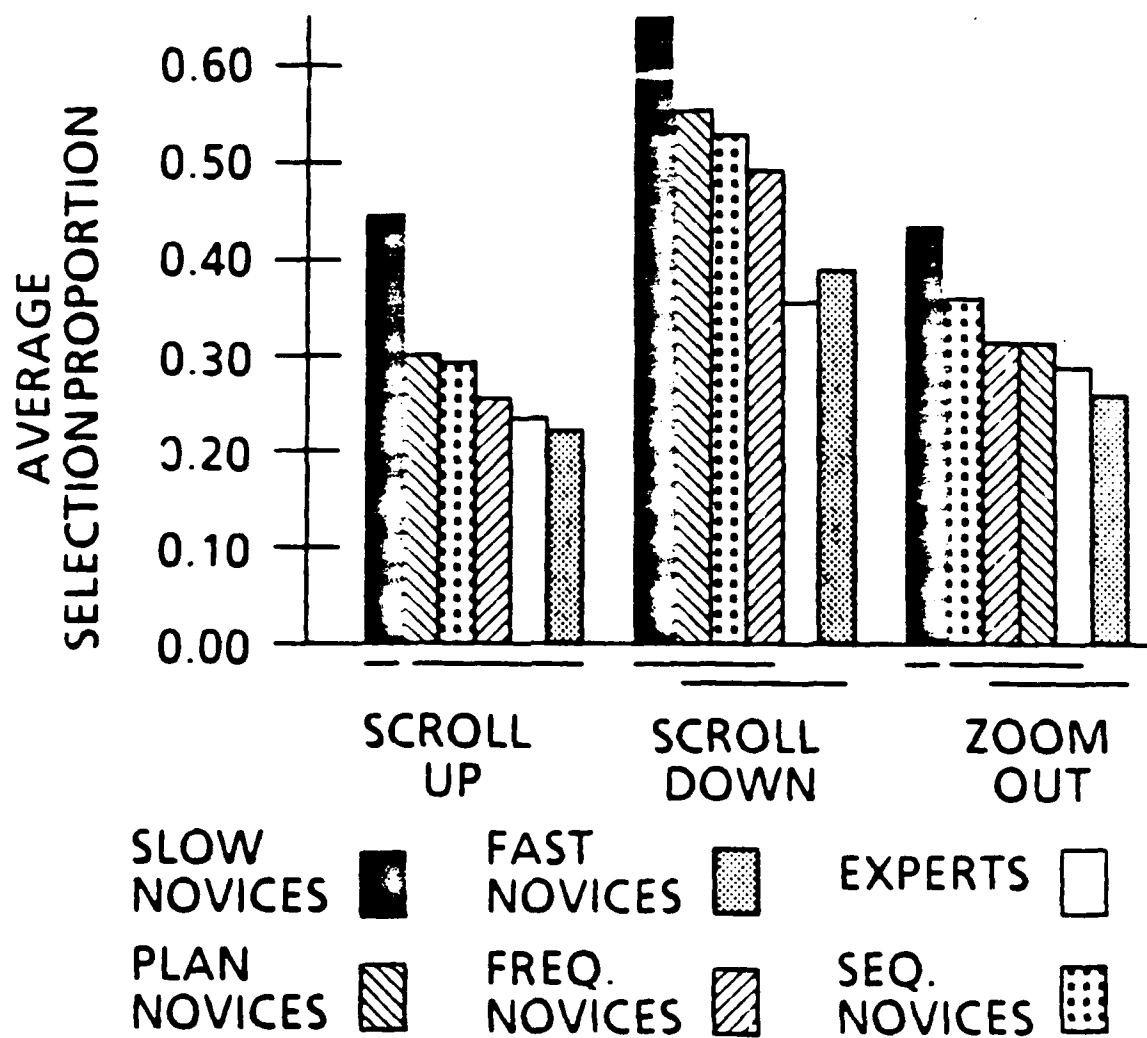
command selection model. For example, the frequency information in combat support files would provide novices the following ranked search procedures (1) FILE SELECT, (2) INDEX, and (3) SEARCH. This command-selection information is useful since the information associated with this problem is embedded in the middle of a file at the second-level of the file hierarchy. Thus, the use of FILE SELECT to position to this file and the use of INDEX and SEARCH to locate file information is fairly sensible advice. In fact, novices may only require this simple feedback on command-selection. In addition, this incomplete model might encourage novices to engage in active learning (4) and elaborative processing (9) which may help them develop effective search strategies.

In the second performance-based model, additional data on command sequence was contained in a second-order transition matrix. This command-sequence model represented the context of command selection by considering the previous command in prescribing the next command to be selected. For example, with the same type of search problem described previously, the most frequently selected search procedures after using FILE SELECT were: (1) INDEX, (2) SEARCH, and (3) SECTION. This information on command transition by experts may help novices learn common sequences of search procedures such as the transition from file selection procedures to procedures for large movements within a file.

The final, cognitive model extended the frequency and sequence models with detailed search plans. These search plans were ordered lists of search procedures with supporting explanations for using search commands. Information used to construct the plans included the frequency and sequence data collected in previous models and verbal comments collected during the final session of the experiment. Plan construction was a complex process described by Elkerton (10) which attempted to represent only the important details of command selection. For example, the

search plan constructed for combat support problems with subgroup information was: (1) FILE SELECT to a combat support file, (2) INDEX to locate subgroup information, and (3) if the information is not located ZOOM OUT and repeat steps 1 and 2. These search plans are promising since command selection is a problem-solving task requiring novices to process a large amount of information on the search problem, the file hierarchy, and how to use search procedures. Representing some of this information in an online aid may decrease the mental processing required when the novice learns to retrieve information.

3. Experiment 2: Evaluation of the Command-Selection Aids

Novice-expert differences in hierarchical file search suggested that an online command-selection aid may improve the skills of slow novices. However, novice heterogeneity poses difficulties when implementing an online aid. For example, some novices may require a large amount of detail in selecting commands and others may only need hints. Indeed, the three command-selection models were developed to determine the required detail for effectively aiding novices. Similarly, novice heterogeneity may pose problems in designing a human-computer dialogue for an online aid. Slow novices may need a computer-initiated dialogue to call attention to their problems, while fast novices may require help rarely and may prefer a user- or mixed-initiative dialogue.

3.1. Method

3.2. Subjects. An additional 36 novices were recruited for the experiment. All novices received the same training and practice as previous subjects to facilitate comparisons between experiments.

3.2.1. Implementation of the command-selection aids. All command-selection models highlighted search procedures on the touch keypad. Frequency and sequence models highlighted and ranked the three search procedures most frequently selected by experts. The plan-based model presented the search plan in the work area of the primary display and highlighted this sequence of search procedures on the touch keypad.

When receiving advice from these command-selection models, novices used either a computer-, user-, or mixed-initiative dialogue. Computer-initiated dialogues only provided advice when a novice selected an infrequently used search procedure. User-initiated dialogues, in contrast, allowed the novice to request command-selection advice at any time through a help key. Finally, in addition to allowing the novice to request advice, the mixed-initiative dialogues displayed a suggestion to press the help key when an infrequently used search procedure was selected.

3.2.2. Experimental design. The command-selection models and dialogue initiatives defined a 3 x 3 mixed-factor design. Therefore, each subject was assigned randomly to a command-selection model and received a different dialogue initiative on Days 3, 4, and 5. The order of these dialogues was completely counterbalanced. A final unaided retrieval session on Day 6 assessed the training effectiveness of the three command-selection aids.

3.3. Results and Discussion

3.3.1. Aided search performance. The results of an unbalanced ANOVA revealed that frequency and plan-based novices were significantly faster (F [5,54] = 8.14, p < 0.05) than slow novices. In fact, aided novices performed at the level of fast novices and experts based on the *post hoc* analysis of the means show in Table 1. Therefore, the

frequency and plan-based aids appeared to improve the search performance of slow novices while not interfering with the performance of fast novices. In contrast, significant performance improvements were not observed for sequence novices during aiding. These results imply that the simplicity of the advice was an important factor for effective aiding. When compared to the frequency and plan-based models, sequence models presented a large amount of constantly changing advice based on a complicated transition matrix.

The relative success of the online aids also was influenced by dialogue initiative. Specifically, a Dialogue Initiative by Presentation Order interaction (F [10,36] = 15.09, p < 0.0001) was uncovered by an ANOVA on search time. Using a Newman-Keuls analysis (α = 0.05), mixed-initiative dialogues were found to be disruptive to novice search performance during the first aiding session (average time of 151 seconds) when compared to the three dialogue initiatives in other presentation orders (average times ranging from 69 to 115 seconds). Based on these results, novices initially may not be ready for the additional decision making to determine whether advice should be presented. This interpretation coincides with data of Carroll and Kay (7) and computer-based training principles (11) which indicate that interactive dialogues can intrude on the coherences of the training task.

3.3.2. Transfer search performance. Despite initial difficulties in using the online advice, the outcome of an unbalanced ANOVA on search times in the final unaided session revealed that previously aided novices were faster at retrieving information (F [5,54] = 7.14, p <0.0001). As illustrated in Figure 3, aided novices retrieved information 60% faster than slow novices. Moreover, aided novices Groups sharing a common vertical line are not significantly different, Newman-Keuls (α = 0.05).

AVERAGE TIME IN SECONDS

100 — 75 — 50 — 25 — 0

SLOW NOVICES ◼ FAST NOVICES ▨ EXPERTS ☐

PLAN NOVICES ▨ FREQ. NOVICES ▨ SEQ. NOVICES ▦
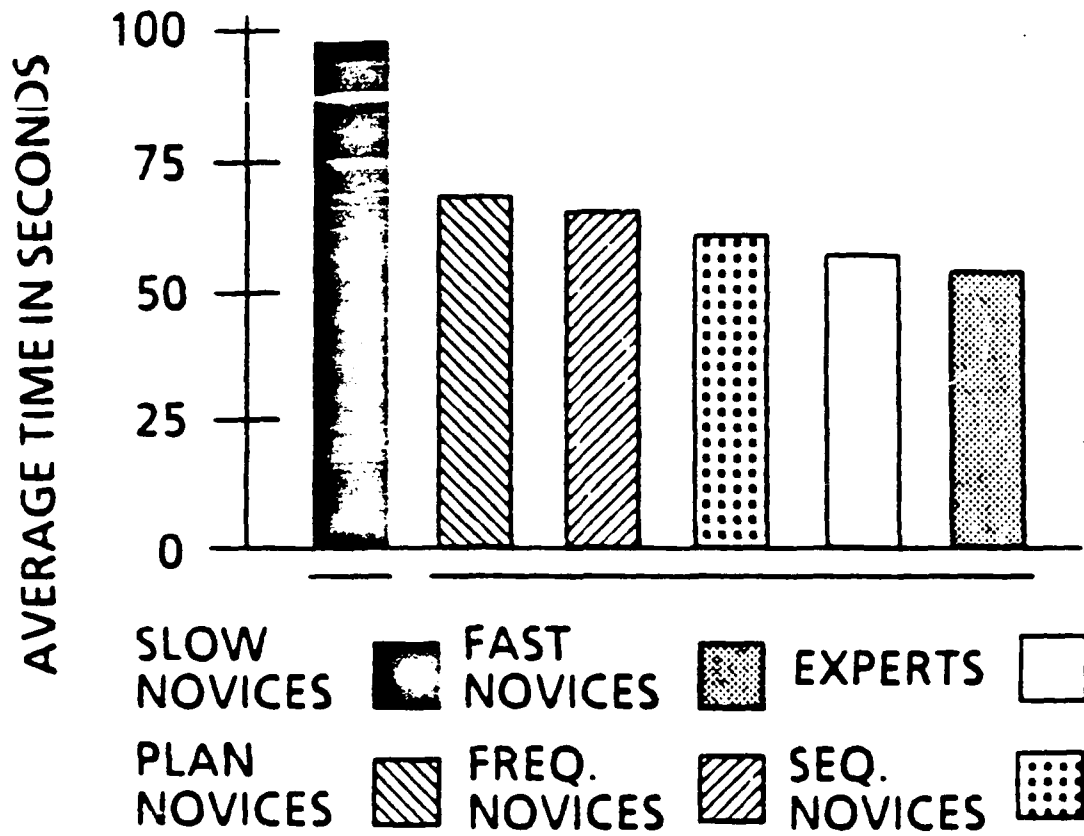
Insert Table 1 Here

----

during transfer retrieved information just as efficiently as the fast novices and experts irrespective of the command-selection advice given during aiding. Therefore, these results indicate that slow novices improved their search performance as a result of previous online aiding.

Unfortunately, the conclusions that can be drawn from this data are limited since slow novices were not explicitly identified in this experiment. However, an analysis of the variability of average search times during transfer for the five subject groups (see Table 2) indicated that the online aids did reduce significantly ($p < 0.05$) the variability in the search performance of aided novices. Furthermore, except for frequency novices, aided novices were similar to the expert group in terms of the variability in search time. Thus, the command-selection aids reduced the heterogeneity of novice search performance.

3.3.3. Search strategy. All command-selection aids facilitated the development of effective search strategies by novices during the experiment. As shown in Figure 4, the use SCROLL UP (F [5,54] = 2.41, $p < 0.05$), SCROLL DOWN (F [5.54] = 2.44, $p < 0.05$), and ZOOM OUT (F [5.54] = 5.10, $p < 0.001$) were reduced significantly when compared to slow novices. Thus, the online aids were effective in eliminating scrolling strategies and emphasizing the structure of the file hierarchy. By providing this consistent practice and feedback, the online aids increased the potential for users learning and automating information retrieval skills (8).

Insert Table 2 Here

TABLE 1 Average Search Times of Aided and Unaided Subjects

Subject Groups	Average Time (seconds)
Slow Novices	119.3
Sequence Novices	103.1
Frequency Novices	94.3
Plan Novices	93.2
Fast Novices	72.7
Experts	67.7

Group sharing a common vertical line are not segnificantly different, Newman-Keuls ($\alpha = 0.05$).

TABLE 2

Comparisons of Variability in Average Search Times During Transfer

Experiment 2 Subject Groups	Experiment 1 Subject Groups	
	Novices $\alpha^2 = 1049.7$	Experts $\alpha^2 = 110.4$
Frequency Novices $\alpha^2 = 338.7$	3.10† (0.05)‡	3.07 (0.05)
Sequence Novices $\alpha^2 = 108.7$	9.66 (0.01)	0.99 (n.s.)
Plan Novices $\alpha^2 = 235.5$	4.46 (0.05)	2.13 (n.s.)

†Main entries are variance rations, F [11,11]. F-ratios are novices/aided novices in the first column and aided novices/experts in the second column.
‡Values in parentheses represent p-values.

3.3.4. Satisfaction with command-selection aids. Subjective responses to the online aids were measured through bipolar adjective scales and ranked preference ratings of the dialogue initiatives. The results from ANOVAs and Newman-Keus procedures ($\alpha = 0.05$) on bipolar responses were consistent with aided performance significantly more pleasing (F [2,33] = 3.37, $p < 0.05$) and dependable (F [2,33] = 4.79, $p < 0.05$) than the sequence models. In terms of dialogue, computer-initiated dialogues were significantly less preferred except with the plan-based models (Friedman's rank sum tests, $p < 0.05$). Thus, plan based models appeared to provide useful advice to the novice, independent of the dialogue.

Insert Figure 4 Here

----

4. General Discussion

The research demonstrated that online command-selection aids can improve novice strategies and performance in file search. In fact, the online aids were particularly effective in decreasing the heterogeneity of novice search performance by eliminating the time-consuming scrolling and ZOOM OUT strategies. This probably resulted from the command-selection feedback which novices could think about and use during this extended experiment.

Consequently, providing command-selection advice is a truly active learning environment which differs substantially from training wheels interfaces (5,6) and scenario machines (7). In these training approaches (5,6,7), the user interfaces were redesigned to reduce the complexity of training so that initial users could explore the word processor with a minimum of errors. Command-selection aids, in contrast, made no attempt to restructure the interface. Novices were presented command-selection advice to provide a stimulating environment for skill acquisition
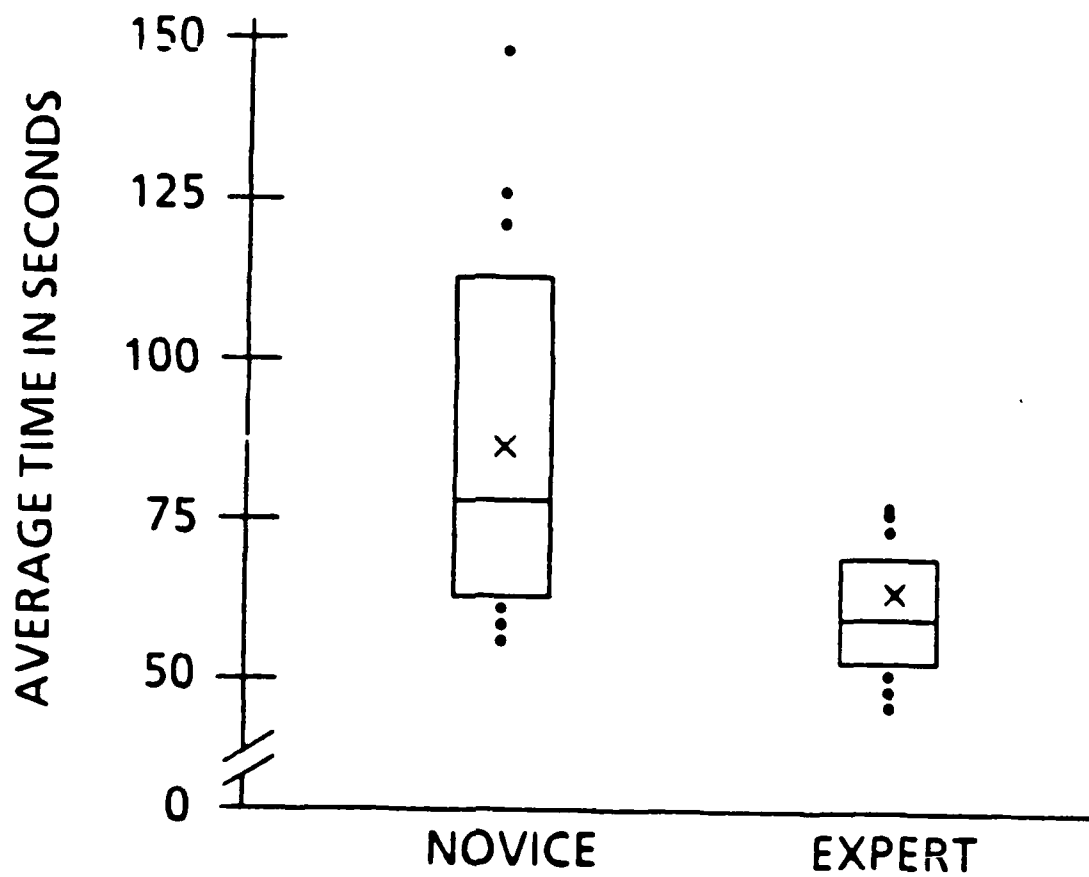
with the full interface assuming that they had basic capabilities to us the advice. Therefore, command-selection aids may help users acquire new strategies during intermediate stages of learning, while training wheels interfaces and scenario machines may be useful during initial learning.

However, the design of an online aid requires that the user be provided appropriate command-selection feedback while minimizing the aiding dialogue presented to the user. As demonstrated, novices provided a mixed-initiative dialogue early in learning may not be ready for the additional decision making required to determine whether they would benefit from the command-selection advice. Similarly, detailed advice focusing on command-sequence may be too complicated for the novice to understand. Therefore, systematic approaches for defining what the user must know to operate an interface should be developed for online command-selection aids. These more formal approaches might use the GOMs models of Card, Moran, and Newell (12) to provide the substantive detail for the command-selection dialogues. Indeed, the subjective benefits offered with these cognitive models could justify the costs associated with model development.

References

(1) Elkerton, J. and Williges, R.C. (1984). Information retrieval strategies in a file-search environment. **Human Factors, 26,** 171-184.

(2) Elkerton, J. and Williges, R.C. (1985). A performance profile methodology for implementing assistance and instruction in a computer-based task. **International Journal of Man-Machine Studies, 23,** 135-151.

(3) Elkerton, J. and Williges R.C. (1984b). The effectiveness of a performance-based assistant in an information retrieval environment. In **Proceedings of the Human**

Factors Society 28th Annual Meeting (pp. 634-638). Santa Monica, CA: Human Factors Society.

(4) Carroll, J.M., and Mack, R.L. (1985). Learning to use a word processor: By doing, by thinking, and by knowing. In J.C. Thomas and M.L. Schneider (Eds.) **Human factors in computer systems** (pp. 13-51). Norwood, NJ: Ablex.

(5) Carroll, J.M. and Carrithers, C. (1984). Blocking learner errors states in a training-wheels system. **Human Factors, 26,** 377-390.

(6) Catrombone, R., and Carroll, J.M. (1987). Learning a word processing system with training wheels and guided exploration. In **Proceedings of CHI + GI 1987** (pp. 169-174) New York: ACM.

(7) Carroll, J.M. and Kay, D.S. (1985). Prompting, feedback and error-correction in the design of a scenario machine. In **Proceedings of CHI'85, Human Factors in Computing Systems** (pp. 149-153). New York: ACM.

(8) Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. **Human Factors, 27,** 285-300.

(9) Reder, L.M., Charney, D.H., and Morgan, K.I. (1986). The role of elaborations in learning a skill from an instructional text. **Memory and Cognition, 14,** 64-78.

(10) Elkerton, J. (1985). **A behavioral evaluation of command-selection aids for inexperienced users** (Tech. Report HCI-85-01). Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

(11) Eberts, R. and Brock, J.F. (1984). Computer applications to instruction. In F.A. Muckler (Ed.), **Human Factors Review: 1984** (pp. 239-284). Santa Monica, CA: Human Factors Society.

(12) Card, S.K., Moran, T.P., and Newell, S. (1983). **The psychology of human-computer interaction.** Hillsdale, NJ: Lawrence Erlbaum.

APPENDIX B

DEFINING SEARCH STRATEGIES IN INFORMATION

RETRIEVAL[1]

Brian C. Hayes and Robert C. Williges

Department of Industrial Engineering and Operations Research,

Virginia Polytechnic Institute and State University,

Blacksburg, Virginia

ABSTRACT

An analysis of user's goals and subgoals in an information retrieval task is undertaken with the aim of determining the behavioral strategies necessary for successful completion of the task. Three such strategies, or plans, were extracted, and were differentiated according to completeness of information provided by the target, the number of commands comprising the plan, and the *sequential nature of command execution.* Analysis of subject performance on targets corresponding to the developed plans revealed significant differences in time to locate target, number of commands used, and the number of different commands. In addition, analysis of command selection frequency indicated that the commands selected for a particular plan were consistent with the commands specified *a priori* as comprising a given plan. The results are discussed in terms of their potential for providing on-line assistance to users.

INTRODUCTION

Task performance is guided by the person's representation of the task, an understanding of the environment in which the task is performed, and an understanding of the tools or methods available with which to perform the task. Card, Moran, and Newell [1] maintain that this task knowledge may be described in terms of the user's understanding of goals, operators, methods, and selection rules (GOMS). This model recognizes that a task is goal-driven and can be described in terms of its constituent subgoals and component tasks, which are hierarchically and sequentially related to each other. Task performance then, becomes a process of identifying the hierarchy of goals and subgoals, and selecting the appropriate methods to fulfill each goal.

Kieras and Polson [2] maintain that the hierarchical goal structure can be extracted from the user's task representation, and can be used to describe concisely the structure of the task from the user's point of view. A variety of methods have been used to extract this information. These methods have included goal-sate analysis [3], planning nets [4], structured sorting procedures [5], and hierarchical cluster analysis [6]. While the methodologies differ considerably, the essential aim of each is to describe task performance in terms of a hierarchy of goals and subgoals necessary to complete the task.

The present study involves the application of a task-analytic approach to determine the goals and component tasks essential to successful performance in an information retrieval task used by Elkerton [5]. The development of plans to describe expert information retrieval behavior based on the abstraction of the goal hierarchy is described, and the validity of these plans is established. The purpose of this study is to demonstrate that a goal-analytic approach to

plan-building is an efficient methodology that yields valid models of expert performance in the information retrieval task to which it is applied.

METHOD

Information Retrieval Task

The information retrieval task required the subjects to locate a particular piece of information (the target), which appeared on one line within a 2,780 line hierarchical file system developed by Elkerton [5]. This file system consisted of 15 files ranging in length from 55 to 447 lines. Each file contained hypothetical textual and numerical information pertaining to army operations, military vehicles, or combat support.

The task environment was implemented on a VAX 11/750 under the VMS operating system. An elevated time-sharing priority was used during the study to ensure consistent response time. Two Digital VT100 terminals were used for the study. One display was used for information retrieval, and the other was used to present the targets. The information retrieval display contained a 7-line window, a line containing the name and length of the current file, an input and message line, a work area, and a touch keypad. Commands were selected through a Carroll Touch Technology touch-entry device and a VT100 (Qwerty) keyboard.

Search Commands

The touch keypad allowed subjects to select from any of the 12 search commands used by Elkerton [5]. These commands included file manipulation commands (FILE SELECT, ZOOM IN, and ZOOM OUT), large movement search commands (SEARCH, SEARCH-AND, SEARCH-NOT, SECTION, and INDEX), and small movement commands (PAGE UP, PAGE DOWN, SCROLL UP, and SCROLL

DOWN). Only a brief description for each of the twelve commands is provided; a detailed description is provided by Elkerton [5].

The file manipulation commands allowed the subjects to move throughout the file hierarchy, whereas all other commands were restricted to the information contained in the current file. The FILE SELECT command was used for direct selection of a lower-level file. Within the top level files in the hierarchy, certain lines were highlighted to indicate that additional information on these selected topics was available in other files. The ZOOM IN command could be used whenever a highlighted line appeared in the window to move to a lower level file in the hierarchy. Conversely, the ZOOM OUT command allowed the subject to exit from the current file and position the top level file in the window.

The large movement commands either allowed for direct string searches (SEARCH, SEARCH-AND, and SEARCH-NOT), or exploited the structure of the current file (SECTION and INDEX). The string search commands allowed the subjects to locate specific alpha-numeric strings in the current file. The SEARCH command permitted a simple string search, and the SEARCH-AND and SEARCH-NOT commands allowed multiple string searches in the 7-line window using Boolean operators. The SECTION command, is analogous to a table of contents in a book, in that it presented a list of the major topics contained in the current file. The window could be positioned to a specific location in the current file by choosing the appropriate topic. The INDEX command presented the subject with an alphabetized list, in the work area of the display, of keywords contained in the current file. The window was positioned to that area of the file containing the chosen keyword.

The small movement commands allowed the subject to scan the current file either seven lines at a time (PAGE UP and PAGE DOWN), or one line at a time (SCROLL UP and SCROLL DOWN).

Development of Plan Models

Elkerton [5] constructed 12 search plans using a four-step process involving a collaboration of information resulting from sorting and clustering procedures, verbal protocols, second-order transition matrices, and path algebra. The plans to be evaluated in the present paper were developed in a less complicated manner. The authors were seeking the minimum number of plans necessary to describe expert subject performance in the information retrieval task. A task analytic approach, similar to that used in the GOMS model of text editing [1] was followed. The approach required the explication of the task in terms of the goals, sub-goals, and sub-task required in order to successfully complete the task.

A review of the targets used by Elkerton [5] led to the hypothesis that the goals and subgoals to be achieved were in large part determined by the amount of information presented in the target itself. Because of the hierarchical nature of the file system, one of the primary goals of task completion is to identify the file in which the target is located. The review of targets revealed three target types, based on the completeness of file identity information. The first type included the *explicit* targets, which essentially revealed the file in which the target was located. For example, "Locate the *T-72* tank with ID code 1234." The second group of targets included the *clued* targets, i.e., a clue as to the proper file location was provided, for example, "Locate the tank that *travels at 50 km/hr* and has ID code 1234." The third type included the *non-clued* targets: those

targets that provided no information as to file identity, e.g., "Locate the tank with ID code 1234."

Once the goal of file identity had been established, the identification of subgoals for task completion followed. According to the GOMS analysis of the task environment, given the available commands, the task would proceed in terms of the following subgoals: position the window to the approximate location in the file (using the large movement commands), and locate the target (using small movement commands). A target of this type might read "Locate the fourth tank in the Sahara regiment." There were, however, very few targets of this type in the Elkerton [5] study (3 of 96 targets), and therefore these targets were not examined in the present study. The great majority of the targets could be located with a simple string search following a file selection procedure; small movement commands were rarely necessary.

Three plans were identified, each corresponding to one of the three target types. a diagram of these plans is presented in Figure 1. Plan A, representing the commands and their sequence for the explicit targets, merely involves a FILE SELECT to the indicated file, and a SEARCH for the indicated information. Plan B outlines the search plan for the clued targets. Subjects first have to determine in which file to locate the target based on a clue given in the target. Thus, the plan became FILE SELECT to an equipment statistics file, SEARCH for the clue, e.g., 50 km/hr, SCROLL UP to the highlighted line indicating the target file, ZOOM IN, and SEARCH for the indicated information. Plan C, for the non-clued targets, consists of a FILE SELECT, SEARCH, and ZOOM OUT command loop until the target was located. In each plan, SEARCH-AND was included with the SEARCH command to account for targets including two or more pieces of information (e.g., "Locate the tank with ID code

1234 *and* 56 battle hours"). Similarly, PAGE UP was included in each plan as an alternative to SCROLL UP.

Insert Figure 1 Here

The accuracy of the plans was initially determined through a data abstraction process. The command selection data for the top 25 percent of the expert subjects in the Elkerton [5] study were examined to determine the exact sequence of command selection. The plans identified above were consistent with expert subject command selection and its sequence.

Target Generation

For each of the three plans, a collection of 36 individual targets was compiled. Several targets were used from the Elkerton [5] study for replication, and some new targets were generated. New targets were developed such that the collection of targets was equivalent across target types in terms of: (1) distribution throughout the file system, (2) information to be input by the user, and (3) possible SEARCH command matches.

Subjects

A total of 12 computer-experienced persons participated in the study. The subjects averaged 3 hours per day of interactive computer use and reported having completed an average of 11 computer courses. Nine of the 12 subjects reported having non-academic computer experience. Subjects were paid a total of $35 for their participation.

Experimental Procedures

The study required 4 days of subject participation. The first day consisted of instruction in the structure and contents of the file system, and in
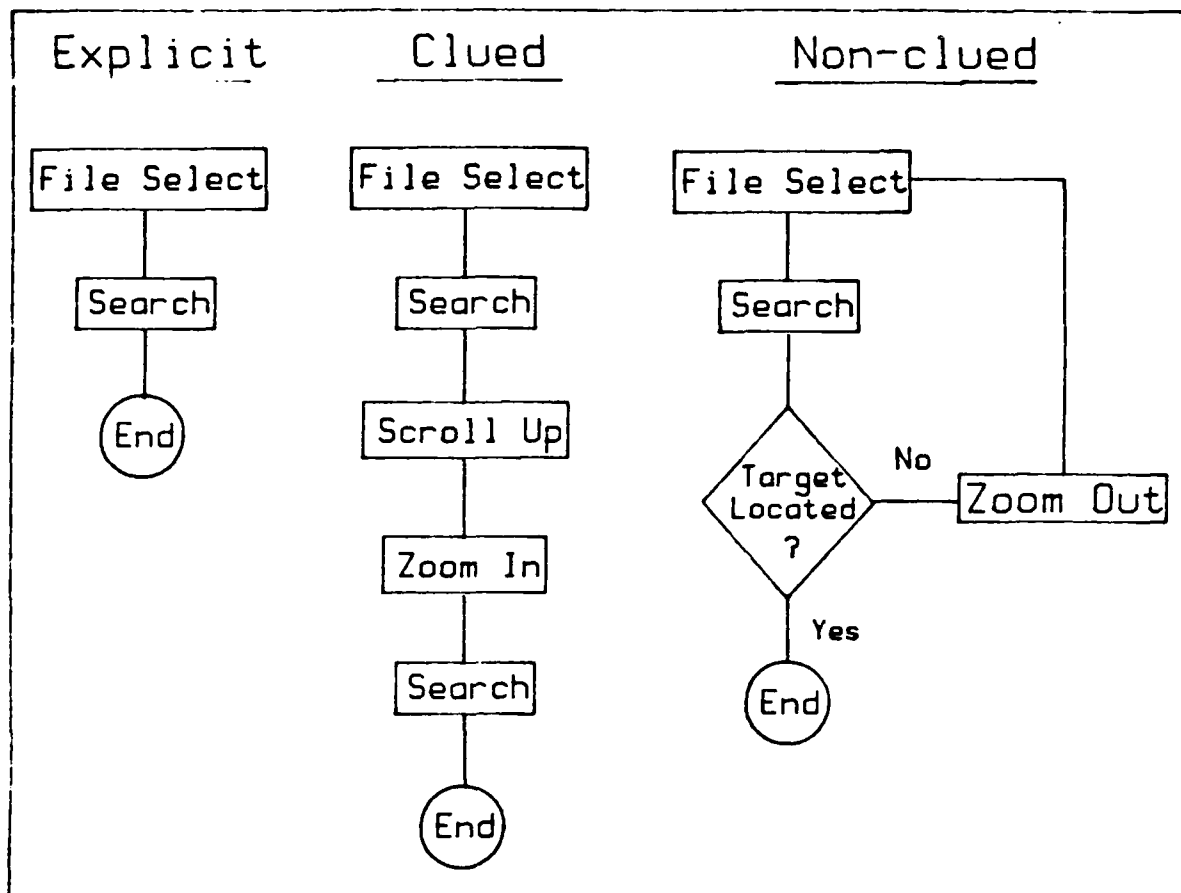
the use of the 12 search commands. This instructional session was presented on a Digital VT100 display and supplemented with oral instruction presented over a Digital DECtalk speech synthesizer. Following the instructional session, the subjects were provided 12 practice trials. A comprehension test of the subjects' knowledge of the file system was administered immediately after the practice trials and a criterion of 70 percent accuracy was required for inclusion in the study. The session lasted approximately two hours.

At the start of each of the three remaining days, subjects were provided four practice trials. Immediately after the practice trials, 3 blocks of 12 trials each were presented, separated by five minute breaks. At the start of a trial, a randomly selected target was presented on the secondary display, and the subject was asked to locate the target in the file system by positioning the window over the location in the appropriate file. A trial was terminated when the subject entered the line number of the target using a function key and digit keys on the VT100 keyboard. The presentation of targets was randomized such that equal numbers of target types were presented each day, and that the presentation of any given target was counterbalanced over the three days of data collection.

RESULTS

Plan-dependent Search Performance

Search performance was analyzed for the following dependent measures: average time per trial seconds, total number of search commands per trial, and the number of different search commands used per trial. Mean values on the dependent measures for the three plan types are presented in Figure 2. Separate analyses of variance (ANOVAs) were conducted on these dependent measures to determine the significance of differences among the three plans.

The three ANOVAs revealed statistically significant differences in time per trial, $F(2,1293) = 144.58$, $p = 0.0001$, total number of commands used per trial, $F(2,1293) = 329.29$, $p = 0.0001$, and number of different commands used per trial, $F(2,1293) = 368.17$, $p = 0.0001$. These results clearly indicate the distinctiveness of the three plans. The values plotted in Figure 2 indicate that the targets in Plan A required the fewest number of commands, the fewest number of different commands, and the least amount of time to locate. Plan C, on the other hand, required fewer different commands than Plan B, yet required the most amount of time to complete. Undoubtedly, this is due to the loop in the plan, i.e., the subject is not given any indication as to what file the target may be located in, and therefore must perform a trial and error file search.

Insert Figure 2 Here

Validation of the Plans

The preceding analysis revealed wide differences in the number of variety of commands selected among the three plans. Therefore, an F-test was conducted to test for an interaction between command selection and target type. The interaction was significant, $F(22,231) = 30.93$, $p = 0.0001$, indicating a significant pattern of plan-dependent command selection. To investigate this finding further, a command selection frequency analysis was conducted, based on a polling procedure developed by Elkerton and Williges [8]. The results of this analysis, delineated according to plan type, are present in Figure 3. The abscissa refers to the proportion of polls; that is, the proportion of trials, averaged across subjects, on which that command was selected a least once. As indicated in Figure 3, the commands selected most frequently by expert
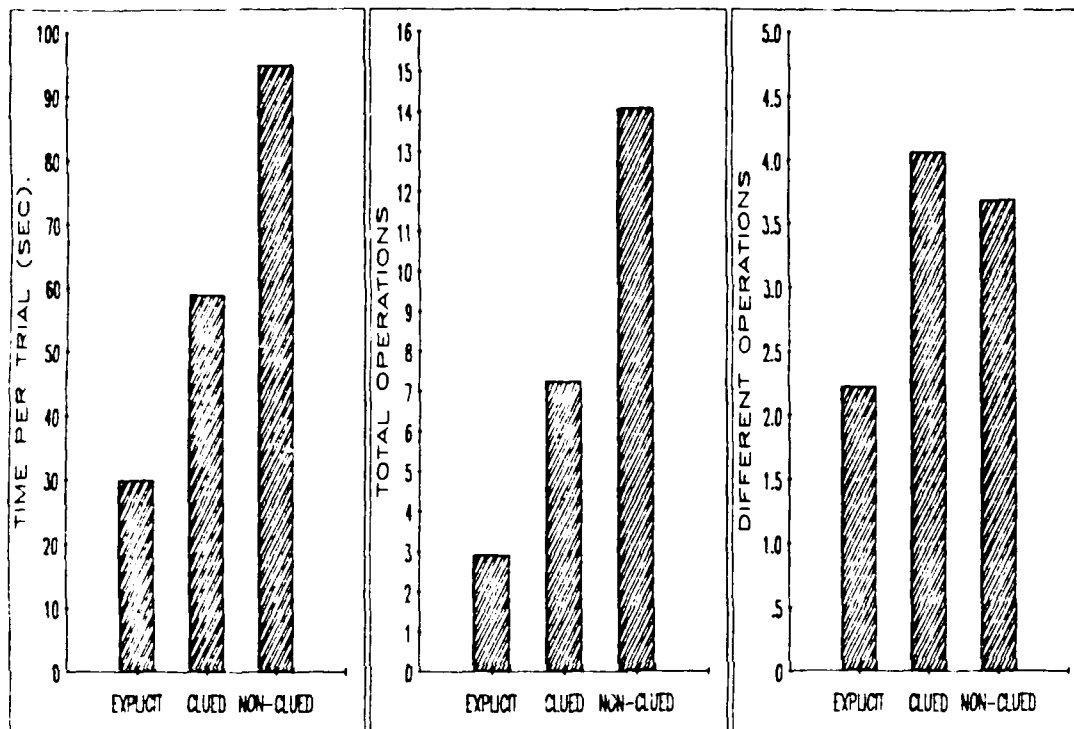
subjects are consistent with the commands specified *a priori* as comprising each of the plans.

Insert Figure 3 Here

DISCUSSION

The results of the present study provide sound support for the goal-analytic approach undertaken, and demonstrate that the *a priori* selection of expert's goals, in terms of the actions to be made, is a valid method for the description of behavior. Furthermore, this description also proved to have predictive validity.

The present study yielded three plans of expert information retrieval behavior. The real utility of these plans, beyond their descriptive properties, may rest in their potential for providing assistance to suers experiencing difficulty in an information retrieval task. Riley and O'Malley [4], for example, concluded that 15 percent of user errors in a text editing task could be attributed to errors in the formulation of goals or plans with which to complete the task. An additional 58 percent of user errors were considered errors in the selection of appropriate commands to achieve their goals. similarly, Elkerton [5] and Hayes, Vicente, and Williges [7] observed that novice subjects selected a larger number and wider variety of commands than computer-experienced subjects. In addition, novice subjects often selected inefficient small movement commands (e.g., scrolling commands) when powerful large movement search commands were available. These results indicate problems with the formulation of plans to direct search behavior and with the selection of commands to
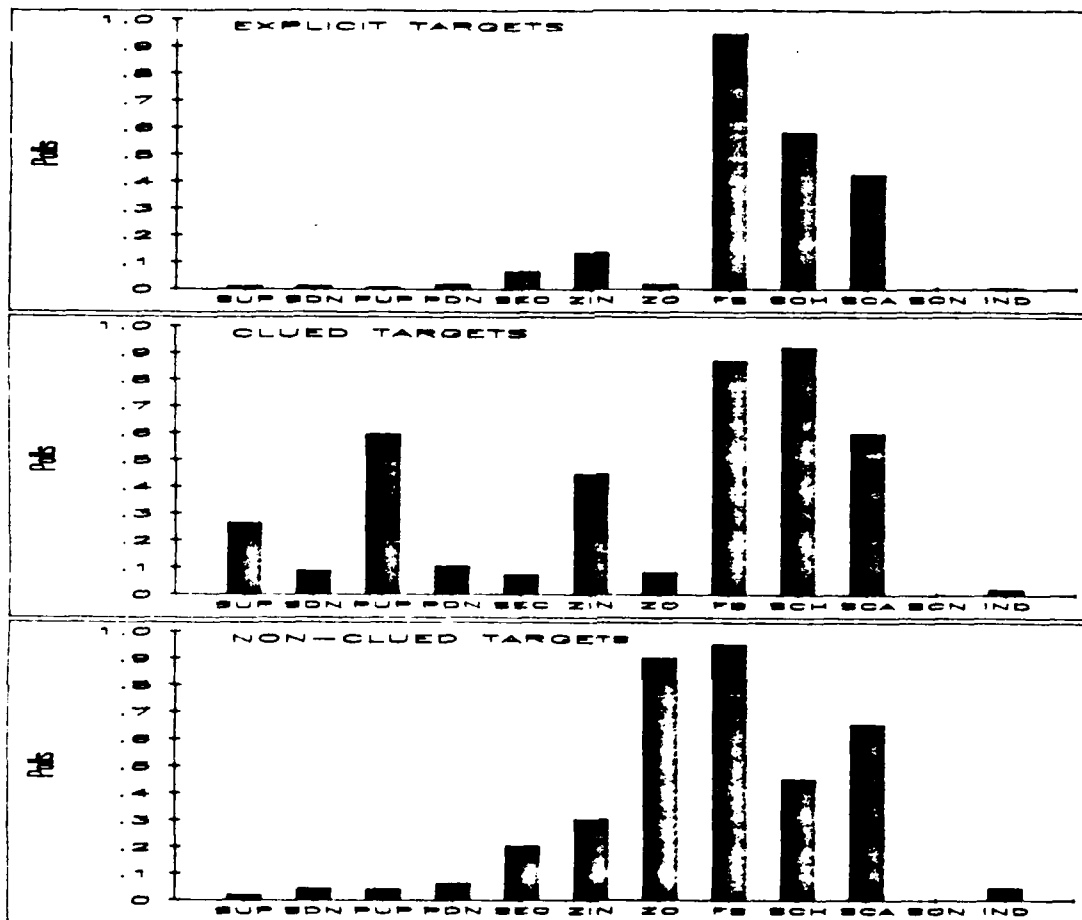
achieve the goals. A plan-based description of expert performance, expressed in terms of plans and thereby improve performance.

The results obtained in the present study involve a rather specific domain, and therefore generalizability is limited. However, the obtained results provide justification for similar undertakings in other domains. Further research is necessary to determine the utility of descriptions of task representation and task performance such that meaningful applications can result.

REFERENCES

[1] S.K. Card, T.P. Moran, and A. Newell, *The Psychology of Human-Computer Interaction*. Hillsdale, New Jersey: Lawrence Erlbaum Associates. 1983.

[2] D. Kieras and P.G. Polson, "An approach to the formal analysis of user complexity," *International Journal of Man-Machine Studies*, Vol. 22, pp. 365-394, 1985.

[3] S.P. Robertson and J.B. Black, "Planning units in text editing behavior," in *Proceedings of the CHI '83 Conference on Human Factors in Computing Systems*, 1983, pp. 217-221.

[4] M.S. Riley and C. O'Malley, "Planning nets: A framework for analyzing user-computer interactions," in *Human-Computer Interaction - Interact '84*, B. Shackel (Ed.). New York: Elsevier Science, 1984, pp. 513-518.

[5] J. Elkerton. "A behavioral evaluation of command-selection aids for inexperienced computer users," Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, December, 1985.

[6] J. Elkerton and B.C. Hayes, "Cluster analysis as a tool for exploring the cognitive factors of a software interface," presented at the Third Annual Midcentral Human Factors/Ergonomics Conference, Oxford, Ohio, June 18-20, 1986.

[7] B.C. Hayes, K.J. Vicente, and R.C. Williges, "Command selection differences among users in an information retrieval task," in *Proceedings of the 1986 IEEE International Conference on Systems, Man, and Cybernetics*, October, 1986.

[8] J. Elkerton and R.C. Williges, "The effectiveness of a performance-based assistant in an information-retrieval environment," in *Proceedings of the Human Factors Society 28th Annual Meeting*, 1984. pp. 634-638.

APPENDIX C

Assaying and Isolating Individual Differences in Searching a Hierarchical File System[1]

KIM J. VICENTE, BRIAN C. HAYES, and ROBERT C. WILLIGES, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University, Blacksburg, Virginia

Individual differences among users of a hierarchical file system were investigated. Results indicate that psychometric tests of vocabulary and spatial visualization are the best predictors of task performance, accounting for 45% of the variance in the data. The spatial predictor was found to be the most influential. This was dramatically illustrated by the fact that, on the average, subjects with low spatial ability took twice as long to perform the task as those with high spatial ability. Surprisingly, experience alone does not predict task performance. A comparison of the frequency of command usage between subjects with high and low spatial ability revealed that those with low spatial ability were getting lost in the hierarchical file structure. These data have implications for redesigning the software interface so as to accommodate people with low spatial ability.

---

[1] This paper appeared in *Human Factors*.

INTRODUCTION

The rapid increase in availability of computer technology has resulted in a corresponding increase in the number of people who are being required to use computers in their daily activities. As a result, users of computer-based systems differ in terms of education, background, experience, and aptitudes. To ensure that all users are capable of using a system effectively and efficiently, it may be necessary to tailor the computer interface to the individual differences which characterize the diversity of the users.

The study of individual differences has a long history in psychology (McFarlane, 1925; Thurstone, 1938). Initial efforts in the area consisted of using factor analysis to describe the nature of mental abilities (Cooper and Mumaw, 1985). Theoretical research has been directed at cataloging the abilities that underly human intelligence (Guilford, 1967). From a more applied perspective, psychometric testing procedures have been used as a selection tool (Ackerman and Schneider, 1985). The basic procedure consists of administering one or more ability tests to a group, and then selecting the people with the highest test scores for the job. In the area of training, individual differences have been used to identify Aptitude x Treatment interactions that enable instructors to tailor the instructional strategies to learner characteristics (Cronbach and Snow, 1977; Savage, Williges, and Williges, 1982). This study takes a similar approach but applies it to the area of software interface design. Thus, individual differences will be used as a means of tailoring the computer interface to the characteristics of the users of those systems.

This research investigates individual differences among users of a computer-based hierarchical file system. Elkerton (1985) found that some computer novices learned to perform this task very quickly, and actually ap-

proached expert performance, while others had much more difficulty in learning the task. A methodology proposed by Egan and Gomez (1985) has been adopted to try to eliminate, or at least reduce, the sources of individual differences in this task. The methodology consists of three phases: assaying, isolating, and accommodating individual differences. The first step, the assay, involves discovering the important sources of individual differences in task performance. As an example, Egan and Gomez (1985) found that spatial memory and age were the main factors in determining how easy it was for novices to use a text editor. Specifically, older people and people with low spatial memory had more difficulty in performing the task. The second step in the methodology involves isolating the effects of user characteristics in particular task components. The purpose of this step is to identify the specific components of the task that are causing the variability in performance. Finally. in the accommodation phase, the problematic task components identified in the second phase are either changed or eliminated. By following this approach it should be possible to design a job so that a wider number of people are able to acquire the skills necessary to perform it.

The goals of this research are to determine which ability measures are accurate predictors of task performance, and which components of the computer task are causing the most difficulty for the slow subjects. In the terminology of Egan and Gomez (1985), this study is restricted to assaying and isolating the individual differences.

EXPERIMENTAL METHOD

Subjects

A total of 30 subjects, 14 of whom were females, volunteered for the study. Subjects' ages ranged from 18 to 31 years, with a mean of approxi-

mately 21 years. All subjects were paid a total of $35 for their participation in the study.

Experimental Design

Subjects were assigned to one of two groups according to the number of hours of their interactive computer experience. The experience of subjects in the Novice Group ranged from 0 to 20 hours, while the experience of subjects in the Experienced Group ranged from 100 to 1000 hours. There were a total of 15 subjects in each of the two groups.

Task Environment

The task environment used in the present study was developed by Elkerton (1985). The subject's goal was to locate a specific piece of information (the target) in a hierarchical file system. There are a total of 15 files in the three-level hierarchy, as shown in Figure 1. The number of lines in a file, also shown in Figure 1, ranges from 55 to 447, with a total of 2780 lines in the entire hierarchy. All of the files contain information about armored personnel carriers, army operations, combat support, and tanks. The information in the files is structured by the use of hierarchies, tables, paragraphs, and lists. On each trial subjects were required to locate a target that existed on only one line of the system.

Insert Figure 1 Here

To locate the target, subjects were provided with 12 search commands that could be selected via a touch screen display, as shown in Figure 2. The display also showed the name of the file currently being displayed (top left corner of the screen), the number of lines in the current file (top right corner

```
                          ARMY OPERATIONS
                           (1100 lines)

        ┌──────────────┬──────────────────┬──────────────────┐
   TANK           TANK DIVISION     ARMORED PERSONNEL      APC DIVISION
   EQUIPMENT      COMBAT SUPPORT    CARRIER (APC)          COMBAT SUPPORT
   STATISTICS     (96 lines)        EQUIPMENT STATISTICS   (209 lines)
   (180 lines)                      (163 lines)

  T-72 TANK    T-62 TANK   PT-76 TANK    BMP-1 APC   BTR-50P APC   MT-LB APC
  (166 lines)  (83 lines)  (156 lines)   (260 lines) (353 lines)   (447 lines)

       T-64 TANK    T-10 TANK        BTR-60 APC    BTR-152 APC
       (139 lines)  (55 lines)       (256 lines)   (107 lines)
```

C-5

of the screen), and a 7-line window into the current file (top half of screen).
Subjects' commands appeared on the input and message line located in the
center of the screen. The output produced by the search commands was
displayed in the work area in the bottom left corner of the screen. Subjects
selected a command by touching the appropriate box on the lower right hand
portion of the screen.

A complete description of the 12 search commands is provided by
Elkerton (1985). The search commands can be divided into three categories:
file selection commands (FILE SELECT, ZOOM IN, and ZOOM OUT), large
movement commands (INDEX, SEARCH, SEARCH-AND, SEARCH-AND-NOT,
and SECTION), and small movement commands (PAGE UP and DOWN,
SCROLL UP and DOWN).

Insert Figure 2 Here

The file selection commands are used to traverse the file hierarchy.
The ZOOM IN procedure is used to move down to a file that is one level lower
in the hierarchy than the currently selected file. Conversely, the ZOOM OUT
procedure enables subjects to move up one level in the hierarchy. The FILE
SELECT command provides a more direct way of selecting a file. Using it,
subjects can select any file that is lower in the hierarchy than the current file.
Thus, it is possible to go directly from the top level to any file at the lowest
level in the hierarchy. These commands enable subjects to select and view
any file in the hierarchy. All of the remaining commands are used to search
for information. It is important to note that the scope of the search commands
is limited to the currently selected file. That is, it is only possible to search for

```
FILE:   MT-L6 Multipurpose Tracked Vehicle (MTV) Inventory      447 LINES IN FILE

19        Bison MTV Regiment 108
20
21     MTV         Front         Maintenance       Battle       Maintenance
22     ID          Location      Battalion         Hours        Code
23
24     845EF5      East          Green-2           184
25     73FD59      East          Green-2           308
```

SCROLL UP	SCROLL DOWN	PAGE UP	PAGE DOWN
SECTION	SEARCH	SEARCH AND	SEARCH AND NOT
INDEX	ZOOM IN	ZOOM OUT	FILE SELECT

information within the current file. To search for information in other files, subjects must first use one of the three file select procedures discussed here to select the new file; only then can they search for information in that new file.

The large movement commands allowed subjects to take advantage of the structure provided within each file. For instance, the SECTION procedure is analogous to a table of contents. When it is selected, a menu of the sections and subsections contained within the current file is displayed in the work area. The 7-line window is then positioned at the beginning of the selected section. The INDEX procedure is analogous to an index of the topics contained in the current file. Its operation is similar to the SECTION command. An index of the topics covered in the current file is displayed in the work area, and the 7-line window is then placed at the point in the file where the topic of interest is located. The search string procedures enable subjects to locate strings of text within the current file. The SEARCH command allows subjects to locate a specific text string. The SEARCH-AND/SEARCH-AND-NOT variants are merely multiple string searches that use Boolean logic. The SEARCH-AND command enables subjects to search for two strings that are located within at most 7 lines of each other. The SEARCH-AND-NOT command adds the constraint that a third string should not appear near the first two. For example, a search for Tank-AND-Regiment-NOT-Orange would try to find a 7-line segment of the file where both Tank and Regiment appear, but Orange does not.

The fine movement commands enable subjects to move the current file through the window either one line (SCROLL UP or DOWN), or seven lines (PAGE UP or DOWN) at a time. Both of these commands can be used continuously. Once subjects thought they had located the target, they hit a function key and entered the line number of the target. The target had to be in the

7-line window for the response to be accepted. If the subject's response was correct, a message to that effect was displayed, and a new trial was presented. If the subject's response was incorrect, a message to that effect was displayed along with a suggestion to continue searching for the target. The primary performance measure is the time taken to find the target. Secondary performance measures include the total number of search operations used per trial, and the number of different search operations used on each trial.

Apparatus

The hierarchical information retrieval system software was implemented on a VAX 11/750 computer. An elevated time-sharing priority was used during the study to ensure that the system response time was consistent across sessions. Information was displayed on two Digital VT100 terminals. The primary terminal displayed the interface shown in Figure 2, while the secondary terminal displayed the targets. The search commands were selected via a Carroll Technology touch entry screen installed on the primary terminal. Additional inputs were typed on a VT100 keyboard.

Pre-testing Sessions

The first two sessions of the experiment, each lasting for about 1 hour and 45 minutes, consisted of pre-testing. Subjects were measured on the battery of 21 predictors summarized in Table 1. Three demographic variables were included in the test battery: sex, hours of experience with interactive computer systems, and number of computer courses taken. These last two were included to test the effect of previous computer experience and knowledge on task performance.

TABLE 1

Candidate Predictors of Task Performance

Spatial   Demographic

Flexibility of Closure   Sex

Perceptual Speed   Computer Experience

Spatial Orientation   Computer Courses

Spatial Scanning

Spatial Visualization   Cognitive Style

Visual Memory   Abstractness

                                 Field Dependency

Verbal

Vocabulary   Other

Reading Rate   Anxiety

Comprehension   Information Processing Rate

Insert Table 1 Here

Previous research has suggested that cognitive style is an important factor influencing behavior in computer-based tasks (Ambardar, 1984; Benbasat, Dexter, and Masulis, 1981; Shneiderman, 1980). Therefore, two measures of cognitive style were included in the test battery. The Abstract Orientation Scale (O'Connor, 1972) characterizes people as being either abstract or concrete. Concrete thinkers tend to be poor at integrating conceptual data in assessing complex problems. They also tend to organize data into relatively few conceptual dimensions. Abstract thinkers, on the other hand, tend to show a greater sensitivity to available cues, as well as a greater ability to use these appropriately and completely in problem solving. Overall, people with an abstract cognitive style tend to be better problem solvers. As an example, Hendrick (1979) found that concrete people were slower and more reluctant to change their set in searching for solutions to a problem. It was hypothesized that slow subjects may have a predominantly concrete cognitive style. The second measure of cognitive style included in the present study was the Embedded Figures Test (1971). With this test, subjects are classified as being either field dependent (FD) or field independent (FI). Subjects who are FI experience parts of a visual field as being discrete from the background, while for those who are FD, perception is strongly dominated by the overall organization of the surrounding field. Scores on the EFT have been found to be related to performance on tasks that require perceptual disembedding, i.e. tasks that require the separation of part of the field from the background. It was hypothesized that successful performance on the information retrieval

system requires subjects to disembed the target from the surrounding information in the file. According to this hypothesis, faster subjects will tend to be predominately FI.

Because the information contained in the files is verbal in nature, subjects with greater verbal ability are expected to perform the task more quickly. In fact, other researchers have found verbal ability to be a good predictor of performance in computer-based tasks requiring reading (Egan and Gomez, 1985). Thus, the Nelson-Denny Reading Test (1973) was included as a measure of verbal ability. This test is composed of three separate sections: reading rate, vocabulary, and comprehension. Each of these was included as a candidate predictor.

To complement the measures of verbal ability, several measures of spatial ability were also included in the test battery. All of these were selected from the Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, and Harmon, 1976). Two forms of each of the following five subsets of spatial ability were selected: flexibility of closure (forms 2 and 3), perceptual speed (forms 2 and 3), spatial orientation (forms 1 and 2), spatial scanning (forms 1 and 2), and spatial visualization (forms 1 and 2). These particular tests were selected because they have been found to be accurate measures of spatial ability (Dupree and Wickens, 1982). In addition, a test of visual memory (form 2) was also included in the test battery. Egan and Gomez (1985) found that visual memory was highly correlated with text editing performance. Due to the spatial nature of the file structure, it is conceivable that subjects who could remember the structure would perform the task more quickly.

Anxiety has been suggested as a correlate of performance on computer-based tasks (Shneiderman, 1980; Spielberger, 1977). Consequently,

the State-Trait Anxiety Inventory (1983) was chosen as a candidate predictor. The STAI provides two measures: state anxiety and trait anxiety. A person's momentary level of anxiety is obtained by subtracting the trait anxiety score from the state anxiety score. If this difference is zero, then the subject is in a neutral state; if the difference is positive, the subject is under stress; if the difference is negative, the subject is relaxed. The final measure included in the test battery was the subject's information processing rate. This was calculated by having subjects perform a choice reaction time task with 1, 2, and 3 bits of uncertainty, and then calculating the slope of the Hick's Law function (Wickens, 1984). In total, 21 predictors were selected for the test battery, some of which were multiple forms of the same construct.

Training Session

After the two pre-testing sessions, subjects participated in a self-paced training session of approximately 2.5 hours. The training was presented on a VT100 terminal and was supplemented with oral instructions presented over a Digital DECtalk speech synthesizer. Initially, all subjects were provided with a hard copy of the files and given detailed instructions as to both the type of information and organization of the document. Subjects were then trained in the use of each of the 12 search commands. Following the instruction for each command, subjects were given two trials to practice the exclusive use of that command. Towards the end of the session, subjects performed 12 practice trials with all search commands available. Finally, subjects received a criterion test of their knowledge of the information retrieval system. All subjects were required to score at least 70% in order to be included in the experiment, thereby ensuring that they all posessed a minimum amount of system know-

ledge. The training was the same for all subjects, regardless of their experience.

Test Session

In the final test session, subjects were presented with 4 warmup trials followed by 2 sets of 12 targets. A five minute break was given between sets. All subjects received the same targets, but the order of presentation was randomized for the test targets. The order of presentation of the warmup trials was kept constant. In summary, the entire experimental procedure consisted of 2 pre-testing sessions, a training session, followed by the data collection session. On the average, subjects took approximately 7 hours to complete the experiment.

RESULTS AND DISCUSSION

Assaying the Individual Differences

The analyses described in this section are directed at identifying the predictors of task performance. First, Pearson product-moment correlation coefficients were calculated between all predictors and the three measures of task performance. Out of the 21 predictors, 6 of these were significantly correlated with at least one performance index as shown in Table 2. Two of the predictors, vocabulary and comprehension, are subsets of verbal ability, while the other four are subsets of spatial ability. From Table 2 one can see that the time to find a target is the most sensitive of the performance measures since all 6 predictors were correlated with it.

Insert Table 2 Here

TABLE 2

Correlations Between Predictors and Performance

| | | Performance Index | |
Predictor	Time	Total Commands	Different Commands
Vocabulary	-0.41 *	-0.42 *	-0.34
Comprehension	-0.37 *	-0.35	-0.26
Spatial Scanning (1)	-0.38 *	-0.34	-0.33
Flexibility of Closure	-0.41 *	-0.30	-0.25
Spatial Visualization (1)	-0.47 **	-0.42 *	-0.44 *
Spatial Visualization (2)	-0.57 ***	-0.46 *	-0.46 *

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

These 6 predictors were then put into a stepwise regression equation to determine the best overall prediction equation. Since time was the most sensitive performance index, it was chosen as the dependent variable for the equation. The overall best equation was chosen by a two step process. First, out of all the one variable models, the one maximizing the variance accounted for was selected. A similar process was performed for all the 2 variable models, 3 variable models, and so on up to the single 6 variable model. Out of these 6 models the one with the lowest Mallows' Cp was chosen as the best equation (Draper and Smith, 1981). This equation is shown in Table 3. The equation contained 2 predictors and accounted for 45% of the variance in the data. One of the predictors, vocabulary, is verbal in nature while the other, spatial visualization, is spatial in nature. The fact that the spatial predictor has a larger beta weight and a larger R-squared value indicates that it is the most predictive of the two variables.

Insert Table 3 Here

Intuitively, one would expect that computer experience would also be correlated with task performance. Indeed, this is the case: people with a greater number of hours of interactive computer experience took less time to perform the task, r (30) = -0.34, p = 0.06. However, when spatial visualization ability is partialled out, the correlation between time and experience is negligible, r (30) = -0.15, p > 0.1. This suggests that experience alone does not affect performance. The main reason that experience and time are correlated is that people with more computer experience also tended to have better spatial ability, r (30) = 0.39, p = 0.03. One explanation for this result is that

TABLE 3

Prediction Equation for Time to Find Target

Raw Score Regression Equation

$$T = 516.29 - 1.67V - 16.09S$$

where T = Time

V = Vocabulary

S = Spatial Visualization

	Variable	
	Vocabulary (V)	Spatial Visualization (S)
Normalized Beta Weight	-0.35	-0.54
R-squared	0.12	0.33
$\underline{F}$	6.14	14.07
$\underline{p}$	0.02	0.001

people with low spatial ability stay away from computers, and therefore never become experienced. Alternatively, working with computers could help individuals develop spatial ability. In fact, previous research indicates that scores on psychometric tests of spatial ability can be improved with training (Blade and Watson, 1955; Brinkman, 1966). Based on the data collected it is not possible to test which of these explanations is correct. However, the finding that experience alone does predict performance is interesting nonetheless.

Another purpose of this study was to assess the magnitude of the individual differences in task performance. If these differences are of practical significance, as opposed to mere statistical significance, then it becomes economically feasible to redesign the software interface in order to accommodate the individual differences. If, on the other hand, the magnitude of the differences is small, it is not worth going through the redesign effort since the benefits will be small relative to the cost involved. To assess the magnitude of the differences, subjects were divided into two equal sized groups according to their vocabulary scores, and an ANOVA with time as the dependent variable was performed. The average times to find a target for the High and Low Verbal groups were not significantly different, F (1,28) < 1, p > 0.1. A similar analysis was performed with subjects divided into two groups according to their spatial visualization scores instead. Results indicate that the Low Spatial group took more time to find a target than the High Spatial group, F (1, 28) = 8.26, p = 0.0077. Figure 3 illustrates the dramatic performance difference between the two groups. On the average, subjects with low spatial ability took twice as long to find a target as those with high spatial ability (average times were 192.02 s and 94.53 s, respectively). The average times to find a target ranged from 53.54 s to 513.17 s, almost one order of magnitude.

The chart shows a bar graph with "Time (s)" on the y-axis (ranging from 0 to 200) and "Spatial Ability" on the x-axis with two categories: "Low" and "High". The "Low" bar reaches approximately 192 seconds, and the "High" bar reaches approximately 95 seconds.
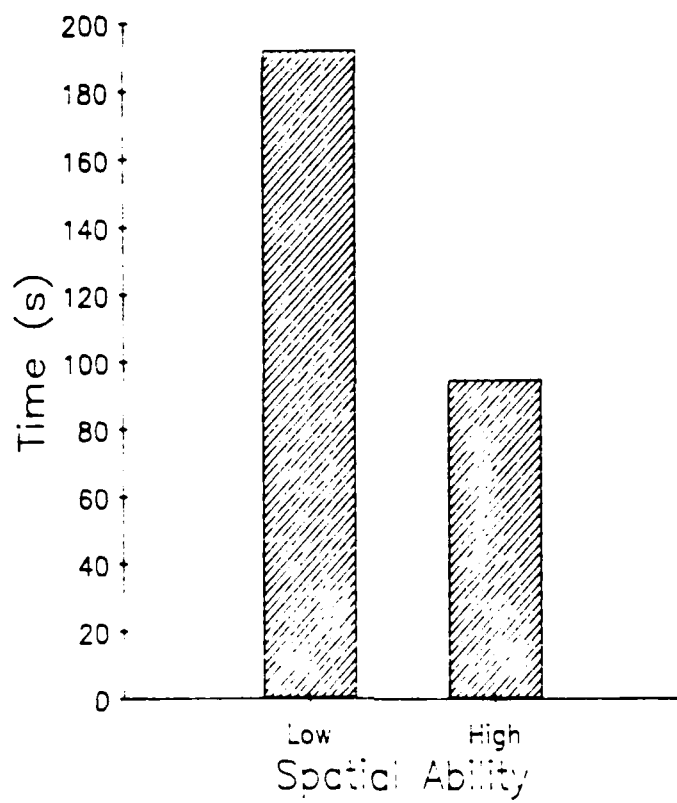
Obviously, redesigning the software interface so as to accommodate people with low spatial ability could result in enormous savings.

Insert Figure 3 Here

Isolating the Individual Differences

Before redesigning the interface, one must first identify which components of the task are causing the most difficulty for the subjects with low spatial ability. By performing this step, one has a much better idea of what parts of the task need to be changed or eliminated in order to accommodate people with low spatial ability. A comparison of command selection strategies of subjects with high and low spatial ability should be useful in isolating the individual differences.

The analysis of search command selection was based on a polling procedure developed by Elkerton and Williges (1985). Conceptually, each subject is given a vector of 12 votes, one for each of the 12 search procedures. A vote is cast for a search procedure if it is selected at least once du~ing a trial. The votes are summed across the 24 experimental trials to yield the total polls for each search command. These are then converted into a proportion indicating the relative use of a search command for each subject. The polling procedure protects against the possibility of bias due to the use of highly repeated search procedures in that each command is counted only once per trial per subject.

As before, subjects were divided into two groups according to their spatial visualization scores. ANOVA's were then performed for each of the 12 commands. The two groups differed in their usage of three commands: ZOOM

OUT, F (1,28) = 26.58, p < 0.0001, SCROLL UP, F (1,28) = 4.88, p = 0.0355, and SCROLL DOWN, F (1,28) = 5.40, p = 0.0277. Subjects with low spatial ability used these three commands more frequently than subjects with high spatial ability, as shown in Figure 4. Elkerton (1985) observed a similar pattern of command selection differences between computer novices and experts. The ZOOM OUT command has a unique function in the information retrieval system; it is the only command which allows subjects to move up in the hierarchy. Therefore, the results suggest that subjects with low spatial ability were going into the incorrect files, and then had to go back up the hierarchy to go into the file where the target was located. In effect, subjects were getting lost in the hierarchical structure.

The greater use of the SCROLL UP and SCROLL DOWN commands by subjects with low spatial ability can be interpreted as an indication that they were getting lost within each file as well. Because they did not know where they were within the file, low spatial subjects scrolled up and down in order to find the information that they were looking for. This result replicates an earlier finding in a study requiring subjects to search within a one-file system, where slower subjects were also found to use scrolling procedures more often than faster subjects (Elkerton and Williges, 1985).

Insert Figure 4 Here

CONCLUSIONS

The assay of the individual differences revealed that psychometric tests of vocabulary and spatial visualization were the best predictors of performance on the hierarchical information retrieval system. The spatial pre-

dictor was the most influential of the two, as illustrated by the fact that subjects with low spatial ability took twice as long to perform the task as those with high spatial ability. Contrary to expectations, the results indicated that experience alone does not predict performance. The attempt to isolate the sources of individual differences revealed that subjects with low spatial ability were getting lost, not only in the hierarchy, but within files as well. The magnitude of the performance differences between high and low spatial ability subjects suggests that redesigning the interface to accommodate those with low spatial ability could result in substantial savings.
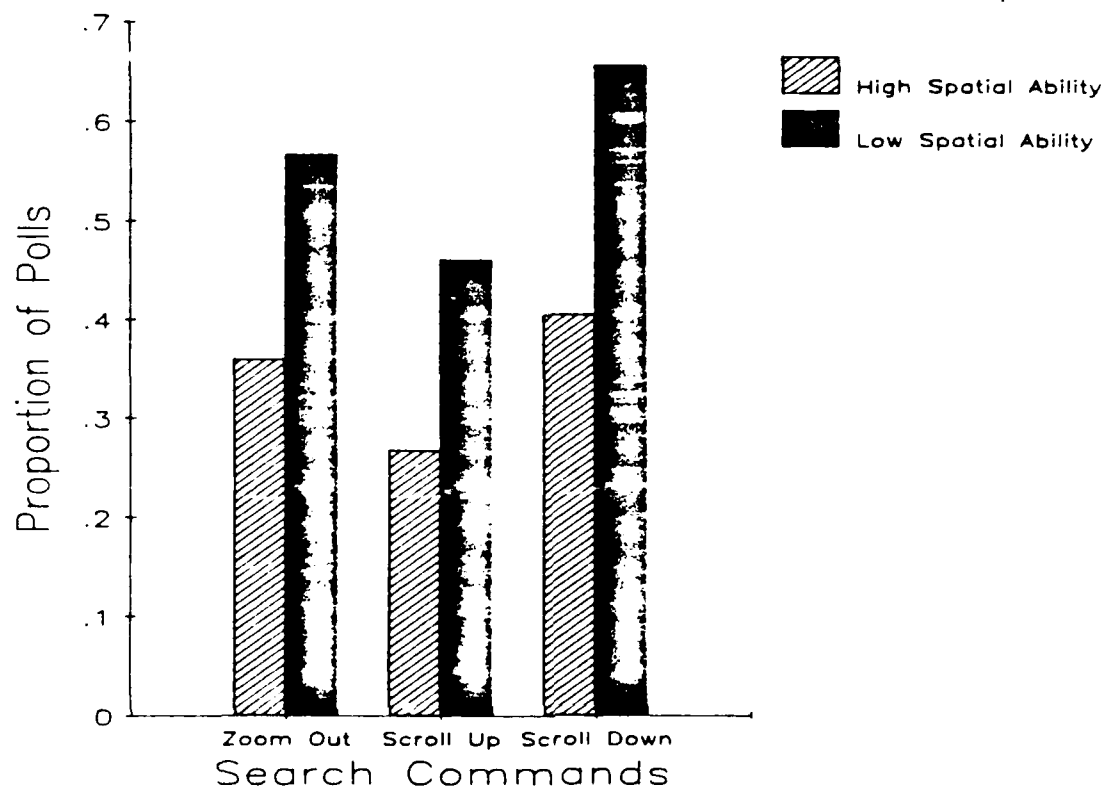
The next step in the Egan and Gomez (985) methodology is to accommodate the individual differences in task performance. This can be accomplished in several ways; by changing the task, by training the users, or by providing on-line support. However, just because the individual differences have been assayed and isolated does not guarantee that the accommodation will be successful. This difficulty is acknowledged by Enan and Gomez (1985, p. 215): "The step of accommodating individual differences not only tests the analyses that precede it, but it also tests the theory of how an experimental manipulation ... will change the original task." Although the assay and the isolation phases locate the locus of the individual differences in specific task components and certain user characteristics, they do not provide the designer with enough information to predict whether or not a given accommodation scheme will be successful.

Future research should be directed toward enhancing the Egan and Gomez (1985) methodology so that the gap between the first two phases and the final accommodation phase is eliminated or reduced. The new methodology should provide a conceptual framework that allows designers to under-

stand the nature of the problem, and thereby select an appropriate accommodation strategy to remedy it. The utility of such a methodology is a direct function of how instrumental it will be in achieving the goal set for this research: To provide empirical evidence that by taking individual differences into account it is possible to design an interface that enables all users, regardless of their abilities, to perform their job effectively and efficiently. Future research will address these important issues.

REFERENCES

Ackerman, P L., and Schneider, W. (1985). Individual differences in automatic and controlled processing. In R. F. Dillon (Ed.), Individual differences in cognition, vol. 2 (pp. 35-66). New York: Academic Press.

Ambardar, A. K. (1984). Human-computer interaction and individual differences. In G. Salvendy (Ed.), Human-computer interaction (pp. 207-211). Amsterdam: Elsevier.

Benbasat, I., Dexter, A. S., and Masulis, P. S. (1981). An experimental study of the human/computer interface. Communications of the ACM, 24, 752-762.

Blade, M. F., and Watson, W. S. (1955). Increase in spatial visualization test scores during engineering study. Psychological Monographs, 69, Whole No. 397.

Brinkmann, E. H. (1966). Programmed instruction as a technique for improving spatial visualization. Journal of Applied Psychology, 50, 179-184.

Cooper, L. A., and Mumaw, R. J. (1985). Spatial aptitude. In R. F. Dillon (Ed.), Individual differences in cognition, vol. 2 (pp. 67-94). New York: Academic Press.

Cronbach, L. J., and Snow, R. E. (1977). Attitudes and instructional methods. New York: Wiley.

Draper, N., and Smith, H. (1981). Appled regression analysis (2nd ed.). New York: Wiley.

Dupree, D. A., and Wickens, C. D. (1982). Individual differences and stimulus discriminability in visual comparison reaction time. In Proceedings of the Human Factors Society 26th Annual Meeting (pp. 809-811). Santa Monica, CA: Human Factors Society.

Egan, D. E., and Gomez, L. M. (1985). Assaying, isolating, and accommodating individual differences in learning a complex skill. In R. F. Dillon (Ed.), Individual differences in cognition, vol. 2 (pp. 173-217). New York: Academic Press.

Ekstrom, R. B., French, J. W., and Harmon, H. H. (1976). Manual for kit of factor-referenced cognitive tests. Princeton: ETS.

Elkerton, J. (1985). A behavioral evaluation of command-selection aids for inexperienced computer users. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Elkerton, J., and Williges, R. C. (1985). A performance profile methodology for implementing assistance and instruction in computer-based tasks. International Journal of Man-Machine Studies, 23, 135-151.

Embedded Figures Test. (1971). Palo Alto: Consulting Psychologists Press.

Guilford, J. P. (1967). The structure of human intelligence. New York: McGraw-Hill.

Hendrick, H. W. (1979). Differences in group-solving behavior and effectiveness as a function of abstractness. Journal of Applied Psychology, 64, 518-525.

McFarlane, M. A. (1925). A study of practical ability. British Journal of Psychology Monograph Supplement, 8.

Nelson-Denny Reading Test, Form C. (1973). Boston: Houghton Mifflin.

O'Connor, J. (1972). Developmental changes in abstractness and moral reasoning. Dissertation Abstracts International, 32, 4109a.

Savage, R. E., Williges, B. H., and Williges, R. C. (1982). Empirical prediction models for training-group assignment. Human Factors, 24, 417-426.

Shneiderman, B. (1980). Software psychology: Human Factors in computer and information systems. Cambridge, MA: Winthrop.

Spielberger, C. D. (1977). Computer-based research on anxiety and Learning: An overview and critique. In J. E. Sieber, H. F. O'Neil, and S. Tobias (Eds.), Anxiety, learning, and instruction. Hillsdale, NJ: LEA.

State-Trait Anxiety Inventory, Form Y. (1983). Palo Alto: Consulting Psychologists Press.

Thurstone, L. L. (1938). Primary mental abilities. Psychometric Monographs, 1.

Wickens, C. D. (1984). Engineering psychology and human performance. Columbus, OH: Merrill Publishing.

APPENDIX D

Accommodating Individual Differences in Searching

a Hierarchical File System[1]

KIM J. VICENTE and ROBERT C. WILLIGES,

Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University, Blacksburg, Virginia

Individual differences among users of a hierarchical file system were investigated. The results of a previous experiment revealed that subjects with low spatial ability were getting lost in the hierarchical file structure. Based on the concept of visual momentum, two changes to the old interface were proposed in an attempt to accommodate the individual differences in task performance. The changes consisted of a partial map of the hierarchy and an analog indicator of current file position. This experiment compared the performance of users with high and low spatial abilities on the old verbal interface and the new graphical interface. The graphical interface resulted in changes in command usage that were consistent with the predictions of the visual momentum analysis. Although these changes in strategy resulted in a performance advantage for the graphical interface, the relative performance difference between high and low spatial groups remained constant across interfaces. However, the new interface did result in a decrease in the within-group variability in performance.

Introduction

The proliferation of new computer systems provides designers with a great deal of flexibility in designing software interfaces. The possibilities are almost endless. However, the mere fact that an interface incorporates the latest technology does not necessarily mean that it is better than previous designs. In fact, it is quite ironic that, while new computer systems (e. g., graphics workstations) provide interface designers with greater flexibility, they can also often result in designs which fail to take into account the capabilities and limitations of their users. Thus, the challenge is to take advantage of the latest technology to build tools that amplify the users' ability to deal with the demands of the task that they are performing.

This research attempts to take advantage of the power of computer graphics to reduce the individual differences in performance among users of a hierarchical file search system. Elkerton (1985) found that some computer novices learned to perform file searching very quickly and actually approached expert performance, while others had much more difficulty in learning the task. The obvious question seems to be: Is there anything that can be done to improve the performance of those that are having difficulty with the task?

The complexity of this question is illustrated by Cronbach and Snow (1977). They cite several studies in the area of instructional strategies that tried to take advantage of subjects' high spatial ability by presenting the instructional material in a pictorial, as opposed to verbal, format. These studies, however, did not show an advantage for the pictorial format. Cronbach and Snow (1977) hypothesize that just because the training programs used diagrams does not make them spatial in nature. Spatial ability

consists of complex cognitive processes such as visualization and orientation that are not necessarily invoked just because the material is presented in a pictorial format. Consequently, the task demands must be matched to the particular characteristics of the students in order to accommodate individual differences.

Conceptual model of accommodation

MODEL

Although the Cronbach and Snow (1977) example comes from the domain of training, the principle generalizes to the present discussion. Any attempt to accommodate individual differences must be based on a conceptual understanding of why certain people are having difficulty in performing the task. The complexities involved are much too great for any approach based on shallow inferences to succeed. Thus, it is useful to think about the major conceptual issues involved in accommodating individual differences. An idealized conceptual model of the problem is presented in Figure 1. The Venn diagram at the top of the figure illustrates the case where there are individual differences in task performance. The large circle represents the capabilities of those subjects who are having difficulty with the task. The smaller circle represents the task demands being imposed on them. It can be seen that the source of the difficulty is that the task is requiring people to do something which they are not capable of doing. Thus, there is a mismatch, or impoverished "cognitive coupling" (Fitter and Sime, 1980), between the task demands and the capabilities of the user, as represented by the shaded portion of the diagram. For people who can perform the task well, this area is negligible.

Insert Figure 1 Here

The model suggests two possible ways of accommodating the individual differences. The first of these, shown in the bottom left corner of Figure 1, is to redesign the task. The idea is to eliminate the mismatch by keeping the task demands within the capabilities of the user. An alternative approach, shown in the bottom right hand corner of Figure 1, is to leave the task as is and to add to the user's capabilities so that the mismatch is eliminated. This can be done either by training the user, or by providing on-line assistance. Regardless of the accommodation strategy chosen, the goal is to eliminate the mismatch between user capabilities and task demands, and thereby increase the cognitive coupling between user and machine. The choice of accommodation strategy for any given situation will depend on the characteristics of the users and task under consideration. For instance, if the source of the mismatch is a fixed ability of the user, it is unlikely that training will remedy the situation. On the other hand, if the source of the mismatch is a subtask that can only be incorporated one way then it will not be possible to redesign the task and yet maintain functionality. Although this conceptual model is idealized and general, it does help one begin to consider the complexities involved in accommodating individual differences. Also, the model may suggest appropriate methodologies for developing an accommodation strategy.

User Capabilities

Task Demands

Mismatch

Accommodation
Process

Redesign
Task

Training or On-line
Assistance

Task Demands

User Capabilities

Task Demands

User Capabilities
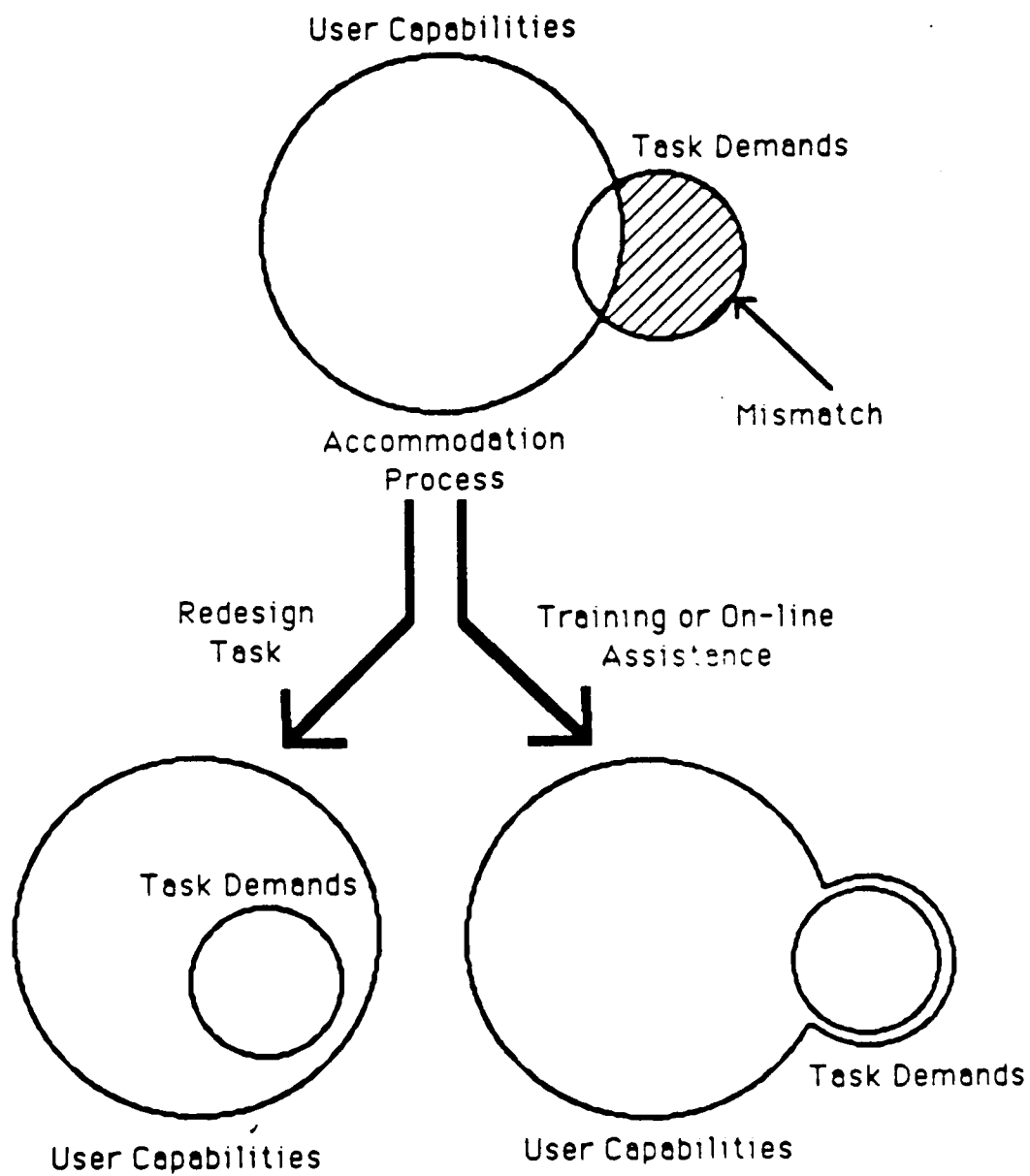
METHODOLOGY

One methodology has been proposed by Egan and Gomez (1985). While not derived from the model described above, their methodology fits in with the concepts described in the conceptual model. The methodology consists of three phases: assaying, isolating, and accommodating individual differences. The first step, the assay, involves discovering the important sources of individual differences in task performance. As an example, Egan and Gomez (1985) found that spatial memory and age were the main factors in determining how easy it was for novices to use a text editor. Specifically, older people and people with low spatial memory had more difficulty in performing the task. The second step in the methodology involves isolating the effects of user characteristics in particular task components. The purpose of this step is to identify the specific components of the task that are causing the variability in performance. Finally, in the accommodation phase, the problematic task components identified in the second phase are either changed or eliminated. By following this approach it should be possible to design a task so that more people are able to acquire the skills necessary to perform it.

It is beneficial to consider how the Egan and Gomez (1985) methodology can be integrated into the model described previously. Figure 2 is a graphical representation of the steps in the methodology couched within the framework of the conceptual model. The problem space is represented as a combination of three domains: that of the user, that of the task, and that of the interface. In addition, the vertical dimension represents a decomposition hierarchy (Rasmussen and Lind, 1981). The whole task, or user, is represented at the top level while lower levels represent parts or subsets of the whole.

Insert Figure 2 Here

Thus, the first step in the methodology, the assay, can be seen as transferring the problem to a lower level in the decomposition hierarchy of the user domain by identifying the relevant subset of user characteristics. Similarly, the second step of isolating the individual differences can be conceptualized as moving to a lower level in the decomposition hierarchy of the task domain so as to identify the problematic task components. These two steps greatly simplify the complexity of the problem by focusing attention on the relevant user and task elements. This reduces the size of the problem space, although developing the accommodation strategy is still not trivial. The problem lies in the fact that it is not possible to compare these two sets of information; one is described in terms of the task and another in terms of the user. What is required at this point is an integration of the two descriptions. This can be done by identifying the demands being imposed on the user by the problematic task components, and the capabilities and limitations of the user from the relevant user characteristics. By bridging the gap between the user domain and the task domain, a better understanding of the reasons for the individual differences in task performance can be acquired. By identifying and understanding the source of the difficulty, it becomes much easier to devise a viable accommodation strategy.

However, there are a few limitations with this approach that merit attention. Egan and Gomez (1985) point out, even when a significant amount of effort has been put into the first two steps, the design remedies are not

User Domain   Interface   Task Domain

User

Task

Relevant User
Characteristics

Problematic Task
Components

Mismatch

User Capabilities
and Limitations

Task Demands

Accommodation
Strategy
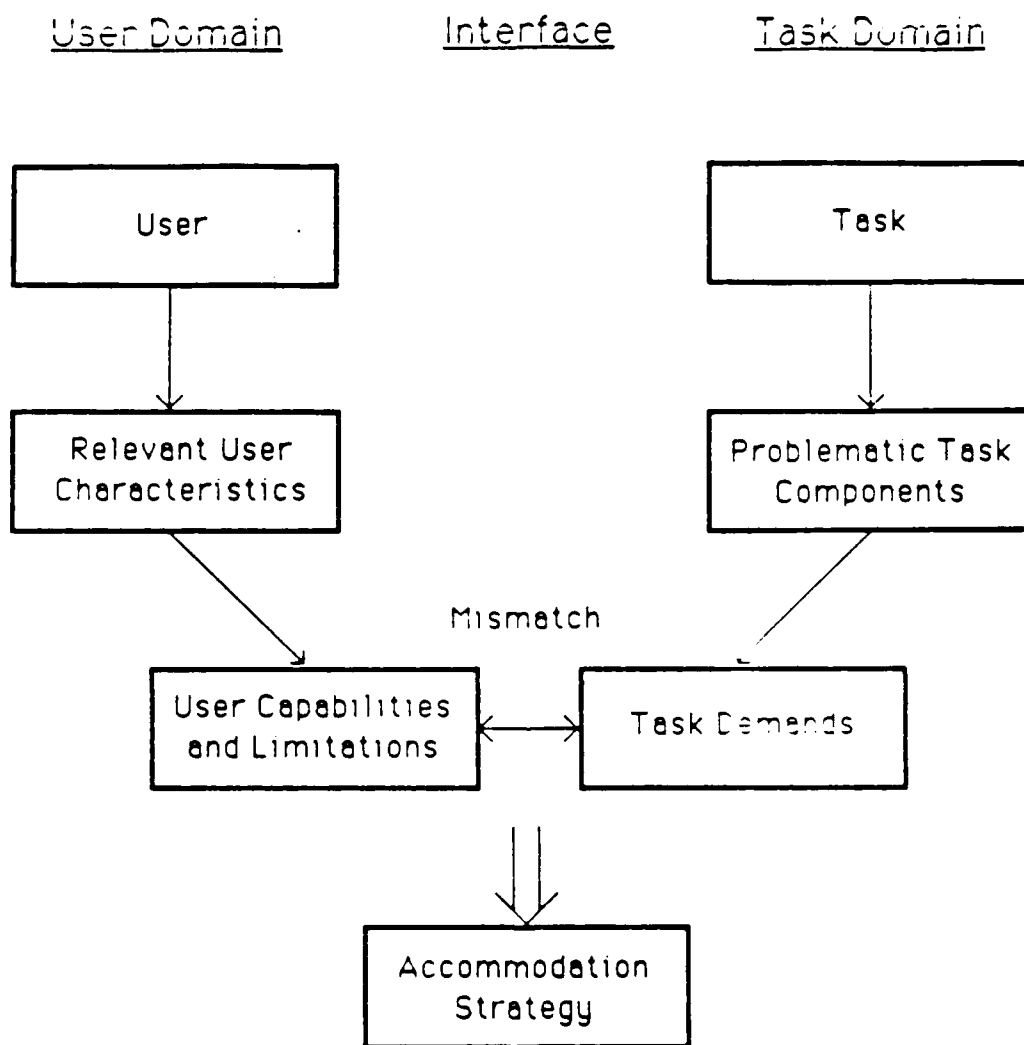
always obvious. Additionally, since the process is one of redesign, it will be biased by the pre-existing design.

Another major difficulty in following the methodology is taking into account the fact that parts of the task may be interrelated in complex ways (Hollnagel, 1983). In some cases, changing or eliminating a task component may have a major effect on how the user perceives the task, and therefore can qualitatively change the way the task is performed. Anticipating such an effect is extremely difficult.

There is also the problem of conflict. In a review of the literature on individualized software interfaces, Rich (1983) observed that system features that make the task easier for one type of user often make it more difficult for another. One resolution to this problem is to design an adaptive interface that changes itself according to the characteristics of the user. As a simple example, a system would be capable of providing extensive and elaborate on-line documentation, or a very brief, terse help function, depending on whether the user was known to be a novice or an expert, respectively.

USING THE METHODOLOGY

Vicente, Hayes, and Williges (in press) dealt with assaying and isolating the individual differences among users of the hierarchical file system. The task consisted of searching for a specific piece of information that existed on only one line of a 2780-line hierarchical file structure. The interface design allowed subjects to view the information contained in the file structure through a 7-line window. Subjects had twelve commands available to them to search for the information they were looking for. The commands can be classified into two classes: those that traverse the file hierarchy (ZOOM OUT, ZOOM IN, FILE

SELECT), and those that search within the currently selected file (SCROLL UP, SCROLL DOWN, PAGE UP, PAGE DOWN, SECTION, INDEX, SEARCH, SEARCH-AND, SEARCH-AND-NOT). The primary performance measure was the time subjects took to find a target.

The Vicente, et al. (in press) assay of the individual differences revealed that psychometric tests of vocabulary and spatial visualization were the best predictors of performance, accounting for 45% of the variance in the data. The spatial predictor was the most influential of the two, as illustrated by the fact that subjects with low spatial ability took twice as long to perform the task as those with high spatial ability (192.02 s compared with 94.53 s, respectively). Surprisingly, the results also indicated that experience alone does not predict performance. The attempt to isolate the sources of individual differences revealed that subjects with low spatial ability used the ZOOM OUT, SCROLL UP, and SCROLL DOWN commands more often than those with high spatial ability. These differences in strategy were interpreted as indicating that subjects with low spatial ability were getting lost in the hierarchy. The results also showed that there was an order of magnitude difference between the fastest and the slowest subject (53.54 s and 513.17 s, respectively). The magnitude of these differences suggests that redesigning the interface to accommodate those with low spatial ability could result in substantial savings.

This study was designed to accommodate the individual differences in task performance found by Vicente, et al. (in press). By examining previous research it should be possible to gain a deeper understanding of the sources of the individual differences and thereby suggest a feasible accommodation strategy.

VISUAL MOMENTUM

Getting-lost in a visual display network has been previously documented by other researchers (Billingsley, 1982; Elm and Woods, 1985: Woods, 1984). Based on research in cognitive psychology. Woods (1984) introduces the concept of visual momentum to account for the getting-lost phenomenon. Visual momentum is a measure of the effort required to integrate and extract information across a set of displays (Woods, 1984). When there is a high degree of visual momentum, it becomes very easy for users to assimilate new data after a transition to a new display. In the case of the hierarchical file system, high visual momentum means that it is relatively easy for users to know where they are in the hierarchy. On the other hand, when visual momentum is low, information extraction is slow and error prone. This results in the getting-lost phenomenon, as indicated by the frequent use of the ZOOM OUT command. Thus, the frequency of usage of the ZOOM OUT command can be used as a measure of visual momentum for the present system.

The frequency of usage of SCROLL UP and SCROLL DOWN are also related to the concept of visual momentum. The difference between these commands and ZOOM OUT is that the latter results in a transition to a new subfile in the hierarchy, while the former results in a transition within a subfile. The frequent usage of ZOOM OUT implies that the user does not have a cognitive map of the hierarchy. The cause behind the frequent usage of the scrolling commands is more subtle. Rather than being a result of poor visual momentum, the frequent use of the scroll commands is a reflection of the subjects' decision to adopt a strategy that inherently has a greater degree of visual momentum associated with it.

Research conducted on the perceptual consequences of various techniques for editing motion pictures explains why this is so. Hochberg (1978) has shown that the time required to integrate successive scenes in a motion picture is related to the amount of overlap in the two scenes. Thus, continuous transitions, such as a pan shot or a tracking shot, contain visual information about the location of one view with respect to the next, and therefore result in high visual momentum. Discontinuous transitions, also called cuts, result in low visual momentum unless the viewer is provided some information with which to understand the transitions. There are several techniques that can be used to improve the visual momentum associated with cuts. One of these is to maximize the amount of overlap between scene cuts; another is to use a long shot to provide a schematic map that allows the viewer to anticipate views (Hochberg, 1978). The time required to comprehend a cut depends on how well the viewer has been prepared to expect the sequence.

The analogy to the hierarchical file system is evident. The scroll commands provide continuous transitions by maximizing the overlap between successive displays, while other commands, such as SEARCH, INDEX, and SECTION provide discontinuous cuts to the next display, i. e., they result in a discrete jump to a new place in the current file. Thus, there is a tradeoff between efficiency and cognitive load. The use of the scroll commands, while very slow, facilitates the integration of information across successive displays by maximizing the overlap between successive views. Commands such as SEARCH, on the other hand, are much more efficient since they take the user to the required information in one step (provided the user is in the correct file and is searching for the correct string); in Hochberg's terms, these commands

provide a discontinuous transition to a new scene. However, these commands put more of a burden on the user to keep track of where he is in the file after each successive move. The problem lies in the fact that, in the hierarchical file system, there is no equivalent of a long shot with which to give the users an overview of the current file.

The data of the Vicente, et al. (in press) experiment indicate that subjects with high spatial ability were better able to keep track of where they were in the file without resorting to the use of the scroll commands. In contrast, subjects with low spatial ability were not able to deal with the added cognitive effort of keeping track of where they were in the file after a discontinuous cut. Thus, they adopted a strategy (i.e., increased use of the scroll commands), which provided them with continuous transitions between views. The drawback was an increase in the average time to find a target. The important point to realize, however, is that these subjects compensated for their lack of spatial ability by using commands which provide them with higher visual momentum than the discontinuous commands. These findings also show that, to attain an acceptable level of performance, people with low spatial ability need displays with more visual momentum than people with high spatial ability.

IMPROVING VISUAL MOMENTUM

Inter-file transitions. The analogy to Hochberg's work suggests various ways in which to improve the visual momentum of the system. For instance, the equivalent of a long shot should result in an improvement in the subjects' ability to determine where they are in the hierarchy, where they can go to next, and where everything else is in relation to their current position. The

computer equivalent of a long shot is a map of the file structure, similar to Figure 3. By making such a map available, the visual momentum of the system should increase. This should be reflected by a decrease in the frequency of use of the ZOOM OUT command, and therefore, a decrease in the average time to find a target. The addition of a map should facilitate the cognitive mapping of the environment. In fact, pictorial representations of complex structures have been found to facilitate the development of a cognitive map (Billingsley, 1982).

However, as shown in Figure 2, to accommodate the individual differences in this particular task it is necessary to aid people with low spatial ability in particular. Thorndyke and Goldin (1981) divided their subjects into two groups, good and poor cognitive mappers, according to the accuracy of their spatial knowledge about their own community. They found that the only characteristic that distinguished between the two groups was spatial ability. In a related study, Goldin and Thorndyke (1981) examined the skills of the same subjects in learning a new environment, map learning, map using, map interpretation, spatial judgements based on a memorized map, and navigation in a new environment based on a memorized map. Their results indicate that good cognitive mappers, when compared to poor cognitive mappers, excel in their ability to encode and retain knowledge of spatial relationships and manipulate their internal knowledge representation in order to compute spatial judgements. More importantly, there was no difference between the two groups' abilities to extract and use information from a map. On the basis of these results, it is expected that providing a map of the hierarchy will aid the performance of subjects with low spatial ability.

Intra-file transitions. In addition to not knowing which file they were in, the slow subjects in the Vicente, et al. (in press) experiment were also having difficulty in keeping track of their position within the currently selected file, as indicated by the excessive use of the scroll commands. This finding is consistent with the research of several cognitive psychologists indicating that cognitive maps are internally represented in a hierarchical structure (Hirtle and Jonides, 1985; Lehtio, Poikonen, and Tuunainen, 1980; McNamara, 1986). These researchers have found that people have multiple internal representations of a given environment, with the higher level representations containing general spatial information and the lower level representations containing more detailed information about the environment. Within this framework, it seems that the slower subjects have difficulty in acquiring and retaining spatial information at both the higher and the lower levels of representation. Presenting subjects with a map should aid them in keeping track of where they are in the hierarchy (higher level representation), but another aid is required for them to be able to know where they are in the current file (lower level representation).

Again, the analogy to film editing suggests a possible remedy. What is needed is a long shot, not of the hierarchy, but of the currently selected file. One possible form that such an aid might take is a rectangular analog indicator of file position. The top and bottom of the rectangle would represent the top and bottom of the current file, and a horizontal line would be drawn in the rectangle to indicate the current position in the file. This type of change to the interface should result in a change in strategy on the part of the low spatial ability subjects. Recall that the hypothesized reason that the low spatials used the scrolling commands so much was that they could not deal with the burden

of keeping track of where they were in the current file after a discontinuous cut. The analog indicator provides a means of relieving this cognitive burden. Thus, there should no longer be any need to tradeoff efficiency for a decrease in cognitive load. The result should be a switch in emphasis from the scroll commands to the more efficient discontinuous commands, such as SEARCH. This change in strategy should lead to a corresponding improvement in performance.

Spatial representations. Will these two proposed changes in the interface improve the performance of both the high and the low spatial ability subjects, or just the low? Other researchers have found that aids of the type proposed above are usually quite effective in improving the performance of all users (Elm and Woods, 1985; Herot, 1984; Sebrechts, Deck, and Black, 1983). Providing information from a user's point of view facilitates the development of a mental model of the system (Nievergelt and Weydert, 1980). Even though the information being provided to the user may be the same, the way that information is represented influences the amount of information that can be processed during problem solving (Mayer, 1976). A spatial metaphor of the type proposed here is especially powerful. By presenting information in this way, users can transfer the knowledge they already possess about how to navigate through an environment, and use it on the problem at hand (Carroll and Thomas, 1982). The use of the spatial metaphor thereby provides a conceptual framework that enables users to organize their knowledge about the system, and allows them to interact with the system in a very natural, almost intuitive, manner. Therefore, it is expected that the proposed aids will improve the performance of all users. Users with low spatial ability are

expected to benefit the most since their overall performance is expected to be poorer. In the case of users with high spatial ability, there is only so much they can improve before the factors limiting their performance become perceptual and psychomotor, rather than cognitive, in nature.

PREDICTIONS

These changes to the interface should help to accommodate individual differences in task performance. The map of the hierarchy should cause a decrease in the frequency of usage of the ZOOM OUT command, indicating an increase in visual momentum. It is also predicted that the addition of the analog indicator will result in a decrease in the frequency of usage of the scroll commands, and an increase in the frequency of usage of more efficient, discontinuous commands such as SEARCH. These changes in strategy should also produce a decrease in the average time to find a target. Although the new interface should improve the performance of all subjects, those with low spatial ability are expected to profit the most from the changes.

Method

SUBJECTS

A total of 75 subjects volunteered for the study. Of these, 28 were excluded due to the pre-test criteria and another 7 were excluded because they failed the criterion test after training. Thus, a total of 40 subjects, 7 of whom were females, participated in the experiment. Subjects' ages ranged from 18 to 27 years, with a mean of approximately 21 years. Each subject was paid for participating in the study.

TASK ENVIRONMENT

The hierarchical file system was developed by Elkerton (1985). The subject's goal was to locate a specific piece of information (the target) in a hierarchical file system. There are a total of 15 files in the three-level hierarchy, as shown in Figure 3. The entire hierarchy was comprised of a total of 2780 lines. All of the files contain information about armored personnel carriers, army operations, combat support, and tanks. The information in the files is structured by the use of hierarchies, tables, paragraphs, and lists. On each trial subjects were required to locate a target that existed on only one line of the system.

Insert Figure 3 Here

First, the Verbal interface used by Vicente et al. (in press) will be described. To locate the target, subjects were provided with 12 search commands that could be selected via a touch screen display, as shown in Figure 4. The display also showed the name of the file currently being displayed (top left corner of the screen), the number of lines in the current file (top right corner of the screen), and a 7-line window into the current file (top half of screen). Subjects' commands appeared on the input and message line located in the center of the screen. The output produced by the search commands was displayed in the work area in the bottom left corner of the screen. Subjects selected a command by touching the appropriate box on the lower right hand portion of the screen.
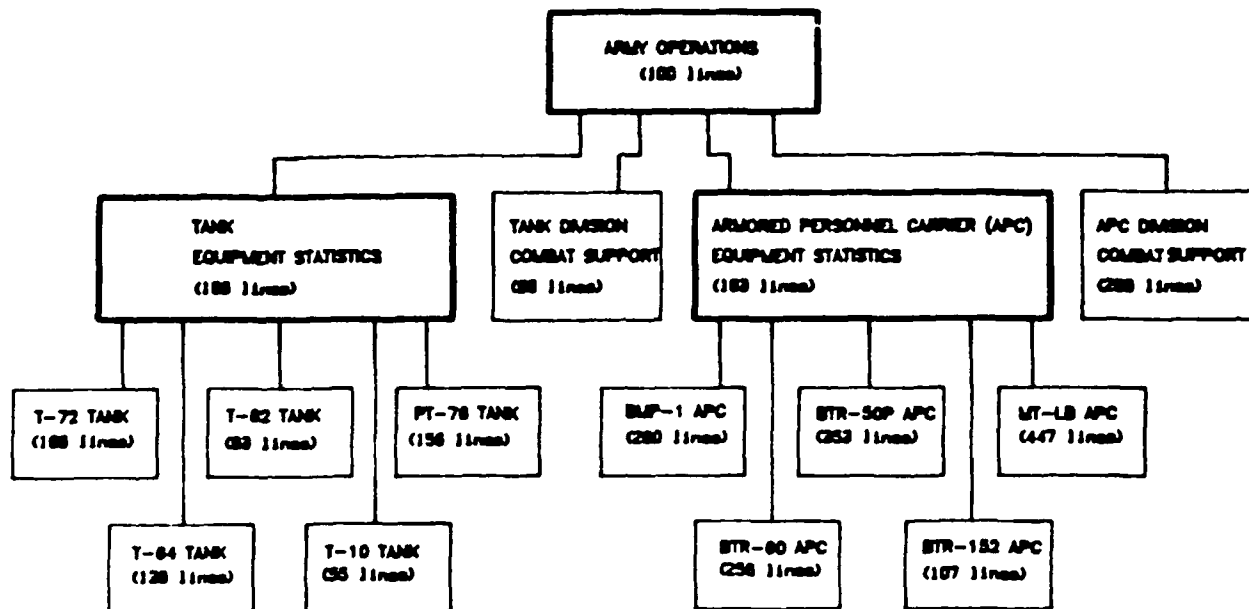
Insert Figure 4 Here

A complete description of the 12 search commands is provided by Elkerton (1985). The commands can be divided into three categories: file selection commands (FILE SELECT, ZOOM IN, and ZOOM OUT), large movement commands (INDEX, SEARCH, SEARCH-AND, SEARCH-AND-NOT, and SECTION), and small movement commands (PAGE UP and DOWN, SCROLL UP and DOWN).

The file selection commands are used to traverse the file hierarchy. The ZOOM IN procedure is used to move down to a file that is one level lower in the hierarchy than the currently selected file. Conversely, the ZOOM OUT procedure enables subjects to move up one level in the hierarchy. The FILE SELECT command provides a more direct way of selecting a file. Using it, subjects can select any file that is lower in the hierarchy than the current file. These commands enable subjects to select and view any file in the hierarchy. All of the remaining commands are used to search for information. It is important to note that the scope of the search commands is limited to the currently selected file.

The large movement commands allowed subjects to take advantage of the structure provided within each file. For instance, the SECTION procedure is analogous to a table of contents. When it is selected, a menu of the sections and subsections contained within the current file is displayed in the work area. The 7-line window is then positioned at the beginning of the selected section. The INDEX procedure is analogous to an index of the topics contained in the current file. Its operation is similar to the SECTION command.

```
FILE:  Army Operations                                    100 LINES IN FILE
┌─────────────────────────────────────────────────────────────────────────┐
│ 1  │            Army Operations                                          │
│ 2  │                                                                      │
│ 3  │ Tank Division                                                        │
│ 4  │                                                                      │
│ 5  │      Tank Equipment Statistics                                       │
│ 6  │                                                                      │
│ 7  │      Tank Division Mission                                           │
└────┴──────────────────────────────────────────────────────────────────────┘
```

SCROLL UP	SCROLL DOWN	PAGE UP	PAGE DOWN
SECTION	SEARCH	SEARCH AND	SEARCH AND NOT
INDEX	ZOOM IN	ZOOM OUT	FILE SELECT

An index of the topics covered in the current file is displayed in the work area, and the 7-line window is then placed at the point in the file where the topic of interest is located. The search string procedures enable subjects to locate strings of text within the current file. The SEARCH-AND/SEARCH-AND-NOT variants are merely multiple string searches that use Boolean logic.

The fine movement commands enable subjects to move the current file through the window either one line (SCROLL UP or DOWN), or seven lines (PAGE UP or DOWN) at a time. Both of these commands can be used continuously. Once subjects thought they had located the target, they hit a function key and entered the line number of the target. The target had to be in the 7-line window for the response to be accepted. If the subject's response was correct, a message to that effect was displayed, and a new trial was presented. If the subject's response was incorrect, a message to that effect was displayed along with a suggestion to continue searching for the target. The primary performance measure is the time taken to find the target. Secondary performance measures include the total number of search operations used per trial, and the number of different search operations used on each trial.

The Graphical interface was similar to the Verbal interface except for the two differences described below. First, the FILE SELECT command, which enables subjects to select any file that is lower in the hierarchy than the current file, was modified. As shown in Figure 5, the Verbal interface presents a list of the selectable files on the primary display, while the Graphical interface presents the files on the secondary display in the form shown in Figure 6. Also, for the Graphical interface a message was printed on the primary display indicating that the files that were displayed in reverse video

in the map had subfiles below them. It is important to note that only the files *below* the currently selected file are shown, just as in the Verbal interface. Of course, when subjects use the FILE SELECT command from either the Tank Equipment Statistics file or the Motorized Rifle Equipment Statistics file, different files than those shown in Figures 5 and 6 are displayed. However, the format of presentation shown in these figures is kept constant for each of the interfaces. Second, instead of listing the number of lines in the current file in the top right-hand corner of the display (see Figure 5), a small rectangular analog indicator illustrated in Figure 7 was used to let the subjects know how long the file is, as well as their current position in the file. These were the only differences between the two interfaces.

Insert Figures 5, 6, and 7 Here

APPARATUS

The hierarchical information retrieval system software was implemented on a VAX 11/750 computer. An elevated time-sharing priority was used during the study to ensure that the system response time was consistent across sessions. Information was displayed on two Digital VT100 terminals. The primary terminal displayed the interface shown in Figure 4, while the secondary terminal displayed the targets. The search commands were selected via a Carroll Technology touch entry screen installed on the primary terminal. Additional inputs were typed on a VT100 keyboard.

FILE: Army Operations	100 LINES IN FILE

1	Army Operations
2	
3	Tank Division
4	
5	Tank Equipment Statistics
6	
7	Tank Division Mission

File number :

FILE   * indicates that subfiles exist

1 * Tank Equipment Statistics	
2 Tank Division Combat Support	
3 * Motorized Rifle Equipment Statistics	
4 Motorized Rifle Division Combat Support	

SCROLL UP	SCROLL DOWN	PAGE UP	PAGE DOWN
SECTION	SEARCH	SEARCH AND	SEARCH AND NOT
INDEX	ZOOM IN	ZOOM OUT	FILE SELECT

```
                    ┌──────────────── Target ────────────────┐
                    │   Locate the T-10 vehicle with the most battle hours   │
                    │                                         │
                    │                                         │
                    └─────────────────────────────────────────┘


                              ┌─────────────────┐
                              │  Army Operations │
                              └─────────────────┘
                                       │
             ┌─────────────┬───────────┴────────────┬──────────────┐
    ┌────────────────┐ ┌───────────────┐ ┌────────────────┐ ┌────────────────────┐
    │    Type 1      │ │    Type 2     │ │    Type 3      │ │     Type 4         │
    │ Tank Equipment │ │ Tank Division │ │ Motorized Rifle│ │ Motorized Rifle Division │
    │  Statistics    │ │Combat Support │ │Equipment Statistics│ │  Combat Support   │
    └────────────────┘ └───────────────┘ └────────────────┘ └────────────────────┘
```

```
FILE:  Army Operations

50  |
51  | Motorized Rifle Division
52  |
53  |      Motorized Rifle Equipment Statistics
54  |
55  |      Motorized Rifle Division Mission
56  |
```

TOP
BOT

SCROLL UP	SCROLL DOWN	PAGE UP	PAGE DOWN
SECTION	SEARCH	SEARCH AND	SEARCH AND NOT
INDEX	ZOOM IN	ZOOM OUT	FILE SELECT

PRE-TESTING SESSION

The first session of the experiment consisted of pre-testing and lasted approximately 30 minutes. During this time, subjects filled out a demographic questionnaire, a test of spatial visualization ability (VZ-2 from Ekstrom, French, and Harmon, 1976), and a test of verbal ability (vocabulary portion of the Nelson-Denny Reading Test, 1973). These were the tests that were found to be the best predictors of task performance in the previous experiment. The spatial ability test was used to screen the subjects and to assign them to either the High or Low Spatial group. This was done by setting cutoff scores for each group, based on the distribution of scores from the previous experiment. The spatial ability test scores from that experiment were submitted to a Kolmogorov-Smirnov goodness-of-fit test to see if they deviated significantly from normality. The analysis indicated that the scores were normally distributed, $D = 0.1226$, $p > 0.20$. Based on the distribution parameters of this sample, cutoff scores were calculated so as to select the upper and lower quartiles of the distribution. Thus, subjects who scored 14 or less out of a maximum 20 points were included in the Low Spatial group, while those who scored 18 or more were assigned to the High Spatial group. Subjects who scored between 15 and 17 were not included in the remainder of the experiment.

EXPERIMENTAL DESIGN

Ten subjects were assigned to the High and Low Spatial ability group based on the pretesting session and were randomly assigned to either the old Verbal interface or the new Graphical interface. In addition, the effects of verbal ability were controlled for through the assignment of subjects to groups. Thus,

the experimental design consisted of a 2 x 2 between-subjects, factorial design with 2 Interface treatments (Verbal and Graphical) and 2 Spatial Ability groups (High and Low Spatial).

TRAINING SESSION

Training subjects on searching the hierarchical filing system was conducted in one session which lasted approximately 2.5 hours. The training for those in the Graphical conditions differed only with respect to the use of the FILE SELECT command and the analog indicator. Otherwise, the two training procedures were identical. The training was presented on a VT100 terminal and was supplemented with oral instructions presented over a Digital DECtalk speech synthesizer. Initially, all subjects were provided with a hard copy of the files and given detailed instructions as to both the type of information and organization of the document. Subjects were then trained in the use of each of the 12 search commands. Following the instruction for each command, subjects were given two trials to practice the exclusive use of that command. Towards the end of the session, subjects performed 12 practice trials with all search commands available. Finally, subjects received a criterion test of their knowledge of the information retrieval system. All subjects were required to score at least 70% in order to be included in the experiment, thereby ensuring that they all possessed a minimum amount of system knowledge. Subjects experiencing the same interface were given the same training, regardless of their spatial ability.

DATA COLLECTION SESSION

The data collection session consisted of four warmup trials followed by two sets of 12 trials on the task. The targets were the same as those used in the previous experiment. To summarize, the entire procedure consisted of one pre-testing session, a training session, followed by a data collection session. On the average, the experiment lasted for approximately 5 hours in total.

Results and discussion

REPLICATION OF RELATIONSHIP BETWEEN EXPERIENCE AND PERFORMANCE

The results of the Vicente et al. (in press) study demonstrated the subjects with more computer experience averaged less time to find the targets. However, when spatial ability was partialled out, there was no correlation between computer experience and task performance. The only reason more experienced subjects were faster was because they tended to have greater spatial ability. In this experiment, again computer experience was significantly correlated with Time, r (40) = -0.34, p < 0.03. When spatial ability was partialled out, however, the correlation between computer experience and task performance was not statistically significant, r (40) = 0.23, p > 0.1. Consequently, these results also indicate that the primary reason that experience is correlated with performance is that more experienced subjects tend to have greater spatial ability.

One explanation for this result is that people with low spatial ability stay away from computers, and therefore never become experienced. This hypothesis of self-selection has been put forth by other researchers as well (Gomez, Egan, and Bowers, 1986). Alternatively, working with computers

could help individuals develop spatial ability. In fact, previous research indicates that scores on psychometric tests of spatial ability can be improved with training (Blade and Watson, 1955; Brinkman, 1966). Based on the data collected it is not possible to test which of these explanations is correct. However, the finding that experience alone does predict performance is interesting nonetheless.

This finding also has important implications for research that attempts to classify users (see Potosnak, 1986). For the most part, studies of this type classify users according to the amount of computer experience they have (Potosnak, 1983). The consistent findings of this research suggest that this may not be a good classification scheme. Instead, researchers should investigate other variables, such as spatial ability in this case, which are the underlying cause of differences in behavior. In situations where relevant user characteristics are not correlated with computer experience, the usual practice of classifying users according to experience will only serve to mask the important underlying relationships between user characteristics and task performance.

EXCLUDED SUBJECTS

It is important to investigate the characteristics of the subjects that were excluded from the experiment due to failing the criterion test of system knowledge. Of the seven subjects who were excluded, six were in the Low Spatial group. The only subject who was in the High Spatial group had an extremely low vocabulary score (27%). Except for one subject, all others had vocabulary scores below 60%. The lone exception had a 95% vocabulary score but scored only 35% on the spatial test, the lowest score of all who were

tested. The pattern seems to be clear. The subjects who were excluded tended to be low in both spatial and verbal ability. The two exceptions scored very high on one test and very low on the other. Finally, the fact that four of the subjects were in the Verbal condition and three were in the Graphical condition suggests that the interface had no effect on subjects' ability to pass the criterion test. An interesting issue for future research would be to determine whether or not these subjects would reach the level of the others, if given enough practice.

THE EFFECTS OF INTERFACE ON PERFORMANCE

The main hypothesis for this experiment was that the Graphical interface would result in improved performance. To test this prediction, ANOVAs were performed with each of the three performance measures as dependent variables. The results are shown in Table 1.

Insert Table 1 Here

For Time, the main effect of spatial ability was significant as expected, $F(1, 36) = 9.32$, $p = 0.004$. Subjects in the High Spatial group took less time on the average to find a target than those in the Low Spatial group (111.1 s compared with 153.0 s, respectively). The main effect of Interface, however, was not significant, $F(1, 36) = 3.40$, $p = 0.07$, but the data were in the predicted direction.

As shown in Table 1, the other two performance measures, average total number of commands used per trial and average number of different commands used per trial, did show significant improvements in performance

TABLE 1

Effects of Interface on Performance

	Interface			
Performance Measure	Verbal	Graphical	F (1, 36)	p
Time	144.7	119.4	3.4	0.07
Total Commands	15.5	12.4	4.0	0.05
Different Commands	4.83	4.32	8.26	0.0067

due to the Graphical interface, $F(1, 36) = 4.0$, $p = 0.05$, and $F(1, 36) = 8.26$, $p = 0.007$, respectively. The subjects in the Graphical groups averaged fewer total commands per trial (12.4 compared with 15.5 for the Verbal groups). The strongest effect of the Graphical interface was in reducing the average number of different commands that subjects used per trial (4.32 compared with 4.83 for the Verbal interface). It was also expected that the Low Spatial group would benefit the most from the Graphical interface, but the lack of significant interactions for all three dependent measures indicates that this did not occur. The High and Low Spatial groups profited equally from the new interface.

The variability in performance also differed across interface groups. The variances in Time for the Verbal and Graphical interfaces were 4231.5 and 1470.72 for the Low Spatials, and 1533.51 and 300.33 for the High Spatials, respectively. Cochran's test for homogeneity of variance (Winer, 1971) was used to test the significance of these differences. Results indicate that the difference in variance for the High Spatials is statistically significant, $C(2, 9) = 0.84$, $p < 0.05$, while the difference for the Low Spatials, although large, is not statistically significant, $C(2, 9) = 0.74$, $p > 0.05$. Thus, performance with the Graphical interface was more consistent across subjects than with the Verbal interface.

Overall, these results demonstrate the superiority of the Graphical interface. The evidence is convincing; compared to the Verbal interface, the Graphical interface resulted in decreases in total number of commands used per trial, different number of commands used per trial, and within-group variability in Time. Contrary to expectations, the improvement in performance was equal across groups of spatial ability, as evidenced by small nonsignificant interactions.

THE EFFECTS OF INTERFACE ON COMMAND USAGE

The concept of visual momentum predicted that the Graphical interface would also result in certain changes in the frequency of usage of certain commands. Consequently, an analysis of subjects' search command selection strategy was conducted. This analysis was based on a polling procedure developed by Elkerton and Williges (1985). Conceptually, each subject is given a vector of 12 votes, one for each of the 12 search procedures. A vote is cast for a search procedure if it is selected at least once during a trial. The votes are summed across the 24 experimental trials to yield the total polls for each search command. These are then converted into a proportion indicating the relative use of a search command for each subject. The polling procedure protects against the possibility of bias due to the use of highly repeated search procedures in that each command is counted only once per trial per subject. The results for those commands which showed a significant main effect of Interface, along with those commands which were predicted to show a significant difference and did not, are given in Table 2.

Insert Table 2 Here

The frequency of usage of the ZOOM OUT command is a measure of how often users get lost in the hierarchy. In other words, it is an indicator of the visual momentum supported by the system. It was predicted that the partial map of the hierarchy would result in an improvement in visual momentum, as indicated by a decrease in ZOOM OUT usage. As shown in Table 2, the data are in the predicted direction but a statistically significant difference was not observed. The lack of significance can be attributed to the

TABLE 2

Effects of Interface on Command Usage

	Interface			
Command	Verbal	Graphical	F (1, 36)	p
ZOOM OUT	0.46	0.42	1.46	0.24
SCROLL DOWN	0.55	0.40	4.44	0.04
SCROLL UP	0.42	0.33	1.63	0.21
SEARCH-AND	0.27	0.35	4.37	0.04
SECTION	0.32	0.13	6.24	0.02

conservative implementation of the Graphical interface. If the map of the hierarchy was displayed all the time, rather than only when the FILE SELECT command was chosen, then there probably would have been a significant decrease in the usage of ZOOM OUT.

The visual momentum analysis also predicted that the analog indicator would cause a decrease in the frequency of use of SCROLL DOWN and SCROLL UP. As shown in Table 2, only the SCROLL DOWN command showed a statistically significant decrease, $F(1,36) = 4.44, p = 0.04$. The fact that the SCROLL DOWN command resulted in a significant difference, whereas the SCROLL UP command did not, can be explained by considering the circumstances under which scrolling commands were typically used in the Verbal interface. Usually, subjects who constantly relied on the scrolling commands would use them right after moving to a new file in the hierarchy. Since the display window is placed at the top of the file after a move to a new file, subjects typically used SCROLL DOWN more often than SCROLL UP. The Graphical interface seems to have been successful in cutting down this excessive use of SCROLL DOWN.

The addition of the analog indicator was also expected to result in an increase in the frequency of usage of more efficient discontinuous commands such as SEARCH, rather than the inefficient scroll commands. As shown in Table 2, this prediction was confirmed by an increase in the usage of the SEARCH-AND command, $F(1, 36) = 4.37, p = 0.04$. This command consists of a search for two or more strings with the Boolean AND operator. The reason why SEARCH-AND showed an increase whereas SEARCH did not is explained by the fact that most targets had several keywords. Thus, it would be more efficient for subjects to use the SEARCH-AND command than the

SEARCH command because it would allow them to take full advantage of the cues provided by the target. The use of SEARCH-AND-NOT was helpful only for a few of the targets. Thus, it is not surprising that the Graphical interface did not result in an increase in the usage of this command. Of course there are other discontinuous commands that the subjects could have chosen to use (e. g. PAGE UP, PAGE DOWN, SECTION, and INDEX). However, the SEARCH-AND command is much m . e accurate and more efficient than any of these other commands, and thus, it was used more often with the Graphical interface.

The decreased use of the SECTION command with the Graphical interface shown in Table 2 was not predicted. At first glance, it is difficult to understand why the Graphical interface would produce this change in command strategy. However, while observing subjects in the Verbal group, it was noted that they would often use the SECTION command when they actually intended to use the FILE SELECT command. A comparison of these two commands, illustrated in Figures 5 and 8, shows why this is so. Both SECTION and FILE SELECT display a list of alternatives that the subjects must choose from. The confusion results from the fact that the format for these two commands is identical. It is easy to see how subjects could get the two commands confused. In the Graphical interface, however, the FILE SELECT command was changed so that it would display a partial map of the hierarchy, as shown in Figure 6, rather than the table shown in Figure 5. This change inadvertently eliminated the confusion that existed in the Verbal interface between SECTION and FILE SELECT. Also, the fact that FILE SELECT was used on almost every trial suggests that the proportion of times that SECTION was mistakenly used in place of FILE SELECT may have been relatively high.

The result was that there was a decrease in the frequency of usage of the SECTION command with the Graphical interface, F (1, 36) = 6.24, p = 0.02. This inadvertent change in strategy provides an excellent example of how difficult it is to anticipate all the consequences of changing or eliminating a task component. As was the case here, these modifications can sometimes change the way the user perceives the task, and therefore can qualitatively change the way the task is performed. Fortunately, in this case the modification resulted in a positive change in strategy.

Insert Figure 8 Here

It was also predicted that the changes in strategy and the resulting improvement in performance would be greatest for the Low Spatial subjects. The lack of significant interactions indicate that this prediction was not confirmed. Thus, the Graphical interface resulted in improved performance for all subjects, but it was not very effective in reducing the individual differences in performance between High and Low Spatial subjects. The only finding indicating a reduction in individual differences was the decreased inter-group performance variability for those subjects receiving the Graphical interface. To explain why the performance changes did not interact with spatial ability, it is useful to reconsider the reasoning that lead to this prediction. It was argued that the High Spatial subjects would not improve as much because their performance was already quite good with the old interface. Thus, it was expected that a ceiling effect would be observed, i. e., the High Spatials would only improve slightly because they would reach a baseline level beyond which they could not improve further. Meanwhile, the Low Spatials, because their

```
┌────────────────────────────────────────────────────────────────────────┐
│ FILE: Tank Division Combat Support                        96 LINES IN FILE │
├───┬────────────────────────────────────────────────────────────────────┤
│ 1 │ Tank Division Combat Support                                         │
│ 2 │                                                                      │
│ 3 │      Artillery Regiment                                              │
│ 4 │                                                                      │
│ 5 │         Artillery Regiment Mission                                   │
│ 6 │                                                                      │
│ 7 │            Provide fire support to the tank divisions making a main advance │
├───┴────────────────────────────────────────────────────────────────────┤
│ File number :                                                            │
├──────────────────────────────────────┬──────────┬──────────┬──────────┬──────────┤
│ FILE   * indicates that subsections exist │ SCROLL │ SCROLL │ PAGE   │ PAGE   │
│                                      │   UP   │  DOWN  │  UP    │  DOWN  │
│  1 *  Artillery Regiment             ├──────────┼──────────┼──────────┼──────────┤
│  2 *  Multiple Rocket Launcher Battalion │        │        │ SEARCH │ SEARCH │
│  3 *  Anti-Aircraft Gun Regiment     │ SECTION │ SEARCH │  AND   │ AND NOT│
│  4 *  Reconnaissance Battalion       ├──────────┼──────────┼──────────┼──────────┤
│                                      │        │ ZOOM   │ ZOOM   │ FILE   │
│                                      │ INDEX  │  IN    │  OUT   │ SELECT │
└──────────────────────────────────────┴──────────┴──────────┴──────────┴──────────┘
```

performance was so poor to begin with, would have much more room for improvement before they experienced a ceiling effect.

Following from this reasoning, there are at least two possible explanations why the Low Spatial subjects did not improve more than the High Spatial subjects. First, it is possible that the duration of the experiment was not sufficiently long enough to enable High Spatials to reach a performance ceiling. Perhaps, if a longitudinal study were conducted, the High Spatials would cease to improve after a certain amount of practice, while the Low Spatials would continue to improve. Second, it is also possible that the manipulation of interface did not reduce the load on spatial ability, but instead reduced the need for some other general ability in which the two subject groups were equal. This would also result in an interface that would be better for all subjects, as was observed. Only more research can determine which of these explanations is correct.

Summary and conclusions

The comparison of the Verbal and Graphical interfaces yielded several interesting findings. Subjects in the Graphical conditions showed a relative decrease in the use of SCROLL DOWN and SECTION, as well as an increase in the use of SEARCH-AND commands. These changes in command usage produced a corresponding improvement in performance for the Graphical interface as compared to the Verbal interface. Except for the unexpected change in SECTION usage, these results are consistent with the predictions of the visual momentum explanation. Contrary to expectations, however, the improvements due to the Graphical interface were the same for High and Low

Spatial subjects. The Graphical interface also resulted in a decrease in performance variability.

To summarize, the new Graphical interface proved to be better than the old Verbal interface, but it was not entirely successful in eliminating the individual differences in performance. Although subjects in the Graphical conditions out performed those in the Verbal conditions, the relative performance difference between High and Low Spatials remained constant across interfaces.

The fact that the new interface did not result in a considerable decrease in individual differences only serves to reinforce the warnings of Egan and Gomez (1985): assaying and isolating the individual differences does not guarantee that a successful accommodation strategy will be developed. The results of this research also reinforce Hollnagel's (1983) statement that very little is known about the effects that changing a task can have on resulting behavior. The inadvertent decrease in the usage of SECTION that was caused by the Graphical interface is a perfect example of a change in a task causing an unexpected change in behavior.

This does not mean that researchers should give up trying to accommodate individual differences. Quite to the contrary, it means that much more research needs to be performed. But for research in this area to have any practical impact at all, a change in focus is imperative. In the past, one of the primary uses of individual differences has been as a selection tool. The approach adopted in this work and the work of Egan and Gomez (1985) is quite different. Instead of trying to find the right person for the task, whenever possible, attempts should be made to design interfaces so that everyone, regardless of their abilities, will be able to use them efficiently and effectively.

In other words, the proposed approach attempts to adapt the task to the person, not the person to the task. Future research in this area should be directed at discovering how to go about achieving this goal.

References

BILLINGSLEY, P. A. (1982). Navigation through hierarchical menu structures: Does it help to have a map? In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 103-107). Santa Monica, CA: Human Factors Society.

BLADE, M. F., and WATSON, W. S. (1955). Increase in spatial visualization test scores during engineering study. *Psychological Monographs, 69*, Whole No. 397.

BRINKMANN, E. H. (1966). Programmed instruction as a technique for improving spatial visualization. *Journal of Applied Psychology, 50*, 179-184.

CARROLL, J. M., and THOMAS J. C. (1982). Metaphor and the cognitive representation of computing systems. *IEEE Transactions on Systems, Man, and Cybernetics, 12*, 107-116.

CRONBACH, L. J., and SNOW, R. E. (1977). *Aptitudes and instructional methods*. New York: Wiley.

EGAN, D. E., and GOMEZ, L. M. (1985). Assaying, isolating, and accommodating individual differences in learning a complex skill. In R. F. Dillon (Ed.), *Individual differences in cognition, vol. 2* (pp. 173-217). New York: Academic Press.

EKSTROM, R. B., FRENCH, J. W., and HARMON, H. H. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton: ETS.

ELKERTON, J. (1985). *A behavioral evaluation of command-selection aids for inexperienced computer users.* Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

ELKERTON, J., and WILLIGES, R. C. (1985). A performance profile methodology for implementing assistance and instruction in computer-based tasks. *International Journal of Man-Machine Studies*, *23*, 135-151.

ELM, W. C., and WOODS, D. D. (1985). Getting lost: A case study in interface design. In *Proceedings of the Human Factors Society 29th Annual Meeting* (pp. 927-931). Santa Monica, CA: Human Factors Society.

FITTER, M. J., and SIME, M. E. (1980). Responsibility and shared decision-making. In H. T. Smith and T. R. G. Green (Eds.), *Human interaction with computers*. London: Academic Press.

GOLDIN, S. E., and THORNDYKE, P. W. (1981, March). *An analysis of cognitive mapping skill* (RAND Note N-1664-ARMY). Santa Monica, CA: The RAND Corp.

GOMEZ, L. M., EGAN, D. E., and BOWERS, C. (1986). Learning to use a text editor: Some learner characteristics that predict success. *Human-Computer Interaction*, *2*, 1-23.

HEROT, C. F. (1984). Graphical user interfaces. In Y. Vassiliou (Ed.), *Human factors and interactive computer systems* (pp. 83-103). Norwood, NJ: Ablex.

HIRTLE, S. C., and JONIDES, J. (1985). Evidence of hierarchies in cognitive maps. *Memory and Cognition*, *13*, 208-217.

HOCHBERG, J. E. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

HOLLNAGEL, E. (1983). What we do not know about man-machine systems. *International Journal of Man-Machine Studies*, *18*, 135-143.

LEHTIO, P. K., POIKONEN, L., and TUUNAINEN, K. (1980). Retrieval of information from a mental map. In R. S. Nickerson (Ed.), *Attention and performance viii* (pp. 381-392). Hillsdale, NJ: LEA.

MAYER, R. E. (1976). Comprehension as affected by structure of problem representation. *Memory and Cognition, 4,* 249-255.

MCNAMARA, T. P. (1986). Mental representations of spatial relations. *Cognitive Psychology, 18,* 87-121.

NELSON-DENNY READING TEST, Form C. (1973). Boston: Houghton Mifflin.

NIEVERGELT, J., and WEYDERT, J. (1980). Sites, modes, and trails: Telling the user of an interactive system where he is, what he can do, and how to get places. In R. Guedj, P. Ten Hagen, F. Hopgood, H. Tucker, and D. Duce (Eds.), *Methodology of interaction* (pp. 327-338). Amsterdam: North-Holland.

POTOSNAK, K. (1986). Classifying users: A hard look at some controversial issues. In *Proceedings of CHI '86 Human Factors in Computing Systems* (pp. 84-88). ACM: New York.

POTOSNAK, K. M. (1983). *Choice of computer interface modes by empirically derived categories of users.* Unpublished doctoral dissertation, The Johns Hopkins University, Baltimore, MD.

RASMUSSEN, J., and LIND, M. (1981). *Coping with complexity* (Riso-M-2293). Roskilde, Denmark: Riso National Laboratory.

RICH, E. (1983). Users are individuals: Individualizing user models. *International Journal of Man-Machine Studies, 18,* 199-214.

SEBRECHTS, M. M., DECK, J. G., and BLACK, J. B. (1983). A diagrammatic approach to computer instruction for the naive user. *Behavior Research Methods and Instrumentation, 15,* 200-207.

SHNEIDERMAN, B. (1980). *Software psychology: Human factors in computer and information systems.* Cambridge, MA: Winthrop. Hillsdale, NJ: LEA.

THORNDYKE, P. W., and GOLDIN, S. E. (1981, March). *Ability differences and cognitive mapping skill* (RAND Note N-1667-ARMY). Santa Monica, CA: The RAND Corp.

VICENTE, K. J., HAYES, B. C., and WILLIGES, R. C. (in press). Assaying and isolating individual differences in searching a hierarchical file system. *Human Factors.*

WINER, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

WOODS, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21,* 229-244.